The Assessment of Risk in Intellectually Disabled Sexual Offenders

JANINE ELIZABETH BLACKER

A thesis submitted to the School of Psychology
The University of Birmingham
Doctorate in Forensic Psychological Practice (ForenPsyD)

Centre for Forensic Psychology
School of Psychology
College of Life and Environmental Sciences
The University of Birmingham
August 2009

Abstract

This thesis explored the application of risk assessment in sexual offenders with intellectual disabilities. A systematic review of the literature appraised the quality and methodology of research examining the predictive validity of risk assessment instruments, this highlighted a lack of research taking into sexual offenders with intellectual disabilities.

The empirical research paper explored the predictive validity of the RRASOR, SVR-20, RM2000-V and the ARMIDILO instruments using a retrospective design on a sample of special needs offenders with intellectual disabilities. Comparisons with mainstream offenders highlighted the difference between the instruments ability to accurately predict risk between the two groups of offenders. The findings suggest that the ARMIDILO can be useful when predicting risk for an intellectual disabled population.

In the next chapter a risk assessment instrument, the RRASOR, was critically reviewed. Following on from this, a case study using an individual approached to risk assessment in an intellectually impaired sexual offender was demonstrated. This chapter emphasises that comprehensive assessment would be a prerequisite to working effectively with offenders with intellectual disabilities in order to address specific intervention needs. A social skills intervention aimed to reduce the level of dynamic risk posed. This chapter also served to outline some of the difficulties associated with risk assessment and management in routine clinical practice. The final chapter concludes by discussing the implications of the findings for clinical practice and offers some directions for future research.

Acknowledgements

The author would like to thank a number of people who have greatly assisted in the completion of this thesis. Professor Anthony Beech, for his continuing supervision and support over the years during the completion of this thesis. Associate Professor Douglas Boer, at the University of Waikato, New Zealand, for his help during my research. The National Probation Service for allowing me to undertake this research. Shihning Chou, for her guidance with the systematic review. Dr Wilcox, and Mary Leyland, for helping in the completion of the risk assessments, and in great appreciation to Sue Hanson and Stella Briggs for all their support over the years. I would like to offer my sincere thanks to all of my friends and family who have been encouraging and supportive throughout the duration of this course, with a special thanks in particular to Jessica Farr and Richard Russell for all their grammatical help.

Table of Contents

List of Tables

List of Figures

Table of Appendices

Chapter One

The Assessment of Risk in Sexual Offenders with Intellectual Disabilities

The assessment of risk is a complex process that involves identifying factors associated with recidivism and predicting the future dangerousness posed by sexual offenders. Accurate risk assessment is crucial in both clinical practice, in decision making and in the identification of appropriate treatment interventions designed to reduce the risk of recidivism (Craig, Beech & Browne, 2006a). The consequences of inaccurate predictions can be detrimental, with great implications for public safety with the potential for further victims. Risk assessment also has significance for the offenders concerned, as it is associated with effective offender management, given that it is influential on decisions relating to length of custodial sentence, selection and treatment planning, release from prison and supervision within the community (Fisher & Thornton, 1993; Craig, Beech & Browne, 2006b; Bengtson & Långström, 2007). The multifactorial nature of sexual aggression makes the assessment of risk an extremely difficult task (Borum, 1996).

There are a number of different methods used to conduct the assessment of risk in sexual offenders. Doren (2002) described six approaches; (1) *unguided judgement*, where case material is examined without any structured assessment or theory prioritizing the relative importance of the data. With this approach the decisions are intuitive or experiential; (2) *guided clinical judgement*, which involves the clinician deriving their own theories which lack support from current research findings and theories of risk. With this method evaluators are more consistent in their assessment procedures across cased than those using unguided procedures; (3) *anamnestic risk assessment,* current and previous contextual and dispositional factors of an offenders' life are considered by the clinician, in order to identify risk factors which may be present and of relevance to the offenders recidivism risk; (4) *research guided clinical judgement*, an a priori group of factors that have been informed by research are used by a clinician as a guide to assessment; (5) *clinically adjusted actuarial approach*, where one or more actuarial instruments are initially used and then adjusted based on clinically derived considerations; and finally (6) *purely actuarial approach*, which is the most mechanical process. The purely actuarial approach involves algorithmic procedures based on mainly historical static risk

factors and coding to arrive at a probability of reconviction. No adjustments are made beyond the instruments results.

The actuarial risk assessment approach is the most widely researched approach and has been found to be superior in comparison to unaided clinical judgement with respect to levels of reliability and accuracy, and more objective with a limited role of discretion (Hanson & Bussière, 1998; Grove, Zald, Lebow, Snitz & Nelson, 2000; Craig, Browne, Hogue & Stringer, 2004). Unaided clinical judgement can be influenced by emotional and nonrelevant considerations and tends to be biased towards the over prediction of risk (Hood, Shute, Feilzer & Wilcox, 2002).

The actuarial approach has an advantage of being a time-effective and easy to code procedure that is grounded on fixed rules. Actuarial instruments for risk assessments are empirically derived from base rates, recorded as recidivism, re-offence or reconviction and provide a probability of future reconviction over a specified follow-up period. Sexual recidivism is any offence-related behaviour, either legal or illegal that has a clear sexual motivation. Sexual reoffending is the perpetration of another illegal sexual act, whether caught or not, whereas sexual reconviction is defined as a subsequent conviction for any sexual offence that occurred after the completion of a sex offender treatment programme (Falshaw, Friendship, & Bates, 2003).

Reconviction rates are ambiguous and unreliable however, as they are known to vary depending upon the age and subgroup of offenders. Recidivism rates fluctuate between studies and also increase with time and length of follow up (Craig, Browne, Stringer & Beech, 2004; Fisher & Thornton, 1993). Base rates also differ greatly according to the definition of sexual recidivism (Falshaw, Friendship & Bates, 2003) and on different setting for example secure hospital or prison populations (Rogers, 2000), as well as the age and subgroups of offenders (Falshaw et al., 2003; Grubin, 1999; Hanson, 1997). When sexual offences go unreported this negatively impacts upon the actuarial data, reducing the reliability and increasing the likelihood of an underestimate of relative risk.

When assessing risk, both static and dynamic factors need to be addressed. Static risk factors are the unchangeable variables such as criminal history, lack of long term relationship and age, and are useful for evaluating long-term risk. Dynamic risk factors on the other hand are those amendable to change such as criminal lifestyle, cognitive distortions or treatment effects (Craig et al., 2004; Craig, Browne, Stringer

& Beech, 2005). Dynamic factors can then be subdivided into *acute* and *stable* factors. Craig et al. (2005) described *acute* dynamic risk factors as those that are rapidly changing and are the contextual factors that signal the onset of offending, such as negative emotional states or substance misuse, whereas s*table* dynamic factors are the relatively persistent characteristics of the offender and have been arranged into four main areas; deviant sexual interest, socio-affective deficits, attitudes supportive of sexual assault and self-regulation problems (Thornton, 2002).

An over reliance on *static* or historical factors in the composition of actuarial tools has been observed, with few measures considering dynamic factors (Beech, Fisher & Thornton, 2003; Craig, Browne & Stringer, 2003; Beech & Ward, 2004; Ward, Polaschek & Beech, 2006). Actuarial risk assessments aim to establish long-term risk, but in doing so they fail to account for acute risk factors that signify imminent risk. Furthermore, by only employing static risk factors, the assessment becomes insensitive to treatment and does not inform or offer guidelines on how to effectively manage the offender in the short term, and reduce the risk of recidivism (Grubin, 1999; Craig, et al, 2005; de Vogel, de Ruiter, van Beek & Mead, 2004), therefore it has been advised that actuarial risk needs to be adjusted based on treatment-related information (Craig et al., 2005). Moreover, static actuarial risk instruments also give no guidance on which psychological factors underlie risk (Ward, Polaschek & Beech, 2006).

In the past decade dynamic factors have gained empirical support. In a meta-analysis, dynamic risk factors such as treatment drop-out were found to be significant predictors of sexual recidivism (Hanson & Bussière, 1998). Similarly, a change in pro-offending attitudes was found to be related to reoffending in child molesters (Hudson, Wales, Barker & Ward, 2002). Furthermore, when Beech, Friendship, Erikson and Hanson (2002) added psychometric measures of dynamic risk (e.g., socio-affective problems) the accuracy of risk prediction was significantly increased beyond the level achieved using only static-factors on the Static-99 (Hanson & Thornton, 2000) actuarial risk assessment, thus demonstrating the importance of including dynamic risk factors into a comprehensive risk protocol. However, whilst stable dynamic risk factors have been theoretically driven, dynamic factors currently lack the empirical support that static factors hold. Stable dynamic risk factors can be used to modify risk based on the levels of psychological problems currently experienced (Ward, Polaschek & Beech, 2006). In the development of actuarial tools,

both stable and dynamic factors have been included within the meta-analytical techniques but dynamic factors have failed to show incremental validity and be predictive of recidivism and so have not been included, despite this dynamic factors are valuable as they provide clinicians vital information on when and where to intervene (Craig, Beech & Browne, 2006a).

The structured professional judgement (also referred to as the guided clinical judgement) approach has been increasingly used in clinical practice. Guidelines to assist in decision making has improved the consistency of decisions but are more time consuming to implement (Hart, Kropp, Laws, Klaver, Logan & Watt, 2003). The Sexual Violence Risk-20 (SVR-20; Boer, Hart, Kropp & Webster, 1997) has been the most evaluated example of the structured professional judgement (Stadtland et al., 2005; Craig, Beech & Browne, 2006b; Craig, Browne, Beech & Stringer, 2006; Dietiker, Dittman & Graf, 2007; Knight & Thornton, 2007). Still too few empirical studies have attempted to evaluate judgements adopting a structured professional guideline approach and with this limited research base it remains difficult to generalise conclusions of the findings (Hart et al., 2003). More recently the Risk for Sexual Violence Protocol (RSVP; Hart et al., 2003) has been evolved from the SVR-20, but has yet to be empirically scrutinised.

The clinically adjusted actuarial approach which adjusts the actuarial results on the basis of clinical judgement therefore helps overcome some of the limitations of purely static actuarial assessments, as current actuarial instruments are not comprehensive in the inclusion of all significant risk and protective factors associated with sexual offending (Doren, 2002). However, this approach is currently being empirically validated and only a few studies to date have tested the utility of the clinically adjusted actuarial methodology. Whilst the clinically adjusted approach allows for the flexibility of clinical adjustment, there is the danger that clinical judgements made will be subject to evaluator bias (Doren, 2002) or may fail to be clearly articulated or not well founded and this may limit the risk scales predictive accuracy (Quinsey, Harris, Rice & Cormier, 1998; Hart, Laws & Kropp, 2003). Given that each different approach to the assessment of risk is not without their problems, Boer (2006) advocates a convergent approach to the assessment of risk, incorporating both types to ensure that all of the important variables from the different approaches are considered. Actuarial instruments would provide a baseline for clinical prognosis,

and the structured professional judgement instruments would ensure that the appropriate variables are used that are related to an individual client's risk.

Intellectually disabled sexual offenders

Our knowledge of sexual offending by men with intellectual disabilities is lagging behind our understanding of mainstream offenders, and it has only recently been recognised as an important area for research (Rose, Jenkins, O'Connor, Jones & Felce, 2002). Currently, there are few recidivism studies on sex offenders with intellectual disabilities and the information relating to sexual reconviction data for this population is therefore much needed as this would inform risk assessment protocols.

The term 'learning disability' (LD) and 'intellectual disability' (ID) are the prevailing professional terms used within the UK to describe individuals whom had previously been labelled 'mentally handicapped'. In other countries the term 'intellectually impaired' is more commonly referred to (Brown & Thompson, 1997). Historically, the term 'mentally handicapped' was used to describe individuals with learning disabilities, however this is no longer used as it is seen to be devaluing (Tudway & Darmoody, 2005). Other terms such as developmental disability, mental retardation, mental subnormality have also been applied over the years, however, for the purpose of consistency the term intellectually disability will be used to throughout the thesis.

To be diagnosed with an intellectual disability in the UK an individual is required to have their Full Scale Intelligence Quotient (IQ) lower than 70, with an age of onset prior to adulthood and have impaired adaptive/social functioning (British Psychological Society, 2001). However, the distinction between those with and those without learning disabilities is ambiguous as a high proportion using learning disability services appear not to have an intellectual disability (Courtney & Rose, 2004). A number of research studies based on offenders with intellectual disabilities have also included those within the borderline intellectual functioning range (Lindsay, for assessing intellectual functioning (Thompson & Brown, 1997).

The term 'special needs', used mainly by the Australian correctional system comprises individuals with mild/borderline intellectual functioning, those with acquired brain injury and those with significant literacy difficulties (Keeling, Rose & Beech, 2006). 'Special needs' is a more flexible description and encompasses a wider range of intellectual abilities. The term 'special needs' was selected for use in the

third chapter of this thesis, as the research sample consisted of borderline functioning as well as intellectually disabled offenders.

Haaven, Little and Petre-Miller (1990) found the recidivism rate for 62 sex offenders with intellectual disabilities was 23%, after completion of a social skills program, whereas Swanson and Garwick (1990) in a sample of low-functioning sexual offenders, found 40% had reoffended after cognitive treatment, although this was based on only 15 individuals who had received group treatment. Klimecki, Jenkinson and Wilson (1994) reported recidivism rates of around 41% for sex related offences based on a sample of 75 offenders with intellectual disability however, this may have been underestimated given that some of their sample had yet to be released from prison. A follow up after the first twelve months of release from prison revealed that 84% had re-offended. In contrast, Lindsay, Elliot and Astell (2004) found that out of 52 sexual offenders with intellectual disabilities, 35% had re-offended or were strongly suspected of re-offending. In a comparison of studies on sexual reconviction rates Craig and Hutchinson (2005) found that the sexual recidivism rate for intellectually disabled sexual offenders was 3.5 times that of non-disabled sexual offenders over a four year follow-up period, with a greater likelihood of reoffending within a shorter time span.

The application of actuarial measures is influenced by base rates, which can only apply to recidivism data that has been detected and such detected rates vary between groups of offenders (Craig & Hutchinson, 2005). Recidivism rates for sexual offenders with intellectual disabilities has been shown to vary and may not be reliable due to the significant under-reporting of the offences in file information, as individuals with learning disabilities are not generally prosecuted for sexual offences and are instead referred to social services or protected by their families (Beech, Fisher & Thornton, 2003; Thompson & Brown, 1997). Such information is of particular interest since the majority of intellectually disabled offenders may not have been subject to the judicial process (Rose et al., 2002). Furthermore, it is the relatively higher functioning intellectually disabled sex offenders that are likely to come into contact with the criminal justice system (Wilcox, 2004).

It questionable as to whether the risk assessment tools originally developed using mainstream sexual offenders populations are as accurate in the prediction of future sexual recidivism with offenders with intellectual disabilities, given that certain risk factor items contained within risk assessment tools are a very prevalent

characteristic of this population such as relationship problems (Hayes, 1991; Day, 1994). Relationship status is an item contained with the Static-99 and RM2000 risk assessment instruments (Thornton, Mann, Webster, Blud, Travers, Friendship & Erikson, 2003), yet intellectually disabled sexual offenders are less likely to have and maintain enduring and lasting relationships, thus automatically placing them at a higher risk level on actuarial measures.

Thesis Aims

The second chapter of the thesis seeks to examine our current knowledge of risk assessment instruments and appraise the research studies on the predictive validity of risk tools, using a systematic approach to investigate the ability of risk assessment tools to predict sexual recidivism in sexual offenders. The findings of these studies and methodological procedures used are then critically analysed.

Traditionally actuarial tools tend to be developed and validated against mainstream sexual offender populations, thus limiting the transferability to other groups of offenders. As yet no static risk assessments have been specifically designed and empirically tested for sexual offenders with intellectual disabilities. To date very little research has been done to establish the predictive accuracy of risk assessment instruments for sexual offenders with intellectual disabilities, and so it remains unclear as to how well risk assessment instruments predict the risk and whether they should be used at all with this population. The third chapter of the thesis therefore aims to bridge this gap by examining the predictive validity of static risk assessment tools applied to this specific population. The SVR-20 (Boer, Hart, Kropp & Webster, 1997), RM2000-V (Thornton et al., 2003), and the RRASOR (Hanson, 1997) risk instruments were compared along with a newly developed dynamic risk instrument, the ARMIDILO (Boer, Tough & Haaven, 2004). The ARMIDILO was designed for sexual offenders with intellectual impairments, and this is the first study to empirically validate this new dynamic risk assessment tool. In order to test the validity of these risk assessment instruments a community based sample of special needs sexual offenders were matched and compared on static risk variables with a sample of non-intellectually disabled offenders. Risk assessment scores and reconviction rates were then compared.

The fourth chapter of the thesis provides an in depth critique of the RRASOR, a static risk assessment tool used within the research study. The reliability, validity and clinical utility of this risk assessment tool are explored.

Chapter five illustrates risk assessment in practice and aims to evaluate the effectiveness of a discrete treatment intervention. Based on an in depth psychological assessment and functional analysis, a social skills intervention was then designed specifically to address a stable dynamic risk factor in a sexual offender with an intellectual disability.

The final chapter in the thesis draws together the main results from within each of the chapters, providing a summary of all the findings and highlighting the limitations of the research conducted, as well as making recommendations for the direction of future empirical research studies into the assessment of risk with intellectually disabled sexual offenders.

Chapter Two

Predictive Validity of Actuarial Risk Assessment Tools to Assess Sexual Offenders for Risk of Reoffending: A Systematic Approach

The assessment of sexual offenders recidivism risk is paramount as the consequences of an inaccurate decision are extremely costly in terms of future victims. It is essential to the judicial system's decision making and has implications for the management and allocation of resources given to the offender (Bengtson & Långström, 2007).
There are a number of different risk assessment tools which have been developed to help clinicians determine the level of risk offenders pose. All risk assessment tools so far have been validated on non-intellectually disabled populations. In the past two decade there has been a huge growth of empirical research into the risk assessment of sexual offenders, as researchers have attempted to address the predictive validity of these tools (Harris & Rice, 2003). In a recent meta-analysis, Hanson and Morton-Bourgon (2007) found the predictive accuracy of actuarial risk measures to be moderate to large. Standards regarding the methodological quality of the research examining risk assessment instruments have not been subject to a comprehensive review.

The aim of this chapter is to examine current research studies on risk assessment tools using a systematic approach to investigate the ability of risk assessment tools to predict sexual recidivism in sexual offenders, and to appraise the methodology employed within these studies. Using the basic principles of a full systematic review, a rapid systematic approach was applied due to the limited time frame. This was accomplished by applying inclusion/exclusion criteria to identify high quality research studies and to appraise the findings accordingly. The quality of studies and methodological limitations were systematically analysed. It should be noted that systematic reviews of this nature are susceptible to publication bias (Togerson, 2006), with the possibility of excluding or overlooking some relevant and important studies, given the limited number of sources searched and the application of strict criteria within this review.

Abstract

Background and Objectives: Assessing the risk posed by sexual offenders is extremely important as the implications of inaccurate assessments hold detrimental consequences. A systematic approach was used to investigate the predictive validity of actuarial risk assessment tools with sexual offenders, and appraise the methodology and quality of the studies in this field.

Search Strategy and Selection Criteria: Electronic databases were searched for relevant studies reporting empirical evaluations of risk assessment tools. The studies identified were subject to inclusion and exclusion criteria, and the quality and clarity of reporting of the studies were then systematically examined with the findings then appraised. The search was restricted to peer-reviewed publications only. Only risk instruments designed specifically to assess risk in adult sexual offenders were included within this review. Narratives, editorials and commentaries were excluded.

Results: An electronic search yielded 7720 hits. 7114 studies were of irrelevant content and a further 43 did not meet the inclusion criteria. Of the remaining publications 10 were excluded due to poor quality and another 10 were not accessible within the time frame. The search resulted in 43 studies being reviewed (40 cohort and 3 case control studies). Within these studies 10 risk assessment instruments were examined. For each instrument the mean AUC values fell within the moderate range of predictive accuracy. Variability of AUC values were reported in regards to age, ethnicity, offender type, treatment status and the length of follow-up period.

Conclusions: This systematic review revealed measurement and selection biases within the studies. Few studies reported on inter-rater reliability or the risk assessors level of training. A general lack of consistency in the types of offences included within the definitions of sexual recidivism was found, and only one study examined sexual offender with intellectual disabilities. Recommendations for future research were made.

Introduction

Risk assessment is crucial in decision making and has great implications for public safety, as the consequences of inaccurate predictions can be detrimental resulting in further victimisation. The assessment of risk posed by sexual offenders serves multiple functions, by determining the intensity and nature of the treatment services and management resources available. Level of risk informs decisions based on length of custodial sentence, selection for treatment and the level of security required whilst in prison, and of supervision in the community prison (Fisher & Thornton, 1993). Researchers in the field of risk prediction have strived towards developing and validating risk assessment instruments for sexual offenders. Not only has there been an increase in the demand for evidence-based risk assessment tools, but there has also been an increase in the literature on the subject within the last decade (Hanson, Morton & Harris, 2003; Harris & Rice, 2003; Craig, Beech & Browne, 2006b).

Approaches to risk assessment

Prior to the 1990's when professionals initially began to assess the recidivism risk of sexual offenders they employed 'unstructured clinical judgements'. Research then suggested the accuracy of this method was little better than chance (Hanson & Bussière, 1998; Hanson et al., 2003). Much energy has since been applied to the development of more structured and accurate procedures for the assessment risk (Hanson & Morton-Bourgon, 2004).

The results of a meta-analysis have shown that empirically derived mechanical actuarial risk assessment instruments have been consistently more accurate in comparison with unguided clinical opinion in the prediction of sexual, violent and general recidivism (Hanson & Morton-Bourgon, 2004). The items selected for these actuarial scales tend to be historical items that have been found to correlate with sexual recidivism in follow-up studies. There are a number of widely used instruments such as the Rapid Risk Assessment for Sex Offence Recidivism (RRASOR; Hanson, 1997); Static-99 (Hanson & Thornton, 2000); Static-2002 (Hanson & Thornton, 2003); Sex Offender Risk Appraisal Guide (SORAG; Quinsey, Harris, Rice, & Cormier, 1998); Minnesota Sex Offender Screening Tool-Revised (MnSOST-R; Epperson, Kaul, Holt, Hesselton, Alexander & Goldman, 1998) and the Risk Matrix 2000 (Thornton, Mann, Webster, Blud, Travers, Friendship & Erikson, 2003).

An alternative approach to the actuarial method is structured professional judgement (also referred to as guided clinical judgement). Decision-making is assisted by professional guidelines reflecting theoretical and empirical knowledge regarding sexual offending and allows clinicians to take into account of both static and dynamic individual-specific factors in the case. Risk factors are drawn from existing empirically researched criteria and selected based on their ability to demonstrate a relationship to future sexual violence. Guidance is provided on how to score those items, and the overall judgement is left to the clinicians' discretion. An example of this approach is the Sexual Violence Risk-20 (SVR-20; Boer, Hart, Kropp & Webster, 1997). This approach has received less empirical evaluation but has been increasingly used in practice.

To evaluate the predictive accuracy of different tools in studies on sexual risk prediction in large cohorts, there are a number of measures available, such as odds ratios, correlations, sensitivity (percentage of reoffenders correctly identified as high-risk on assessment) and specificity (percentages of non-reoffenders correctly identified as low-risk), and Receiver Operating Characteristic (ROC) Area Under the Curve analysis (AUC). An interpretation of the AUC can be seen as the probability that a randomly selected recidivist will score higher on the instrument than a randomly selected non-recidivist. The AUC provides a probabilistic estimates rather than a verbal label to characterise effect sizes. An AUC of .50 represents a chance prediction, whereas an AUC of 1.0 represents a perfect positive prediction. The ROC analysis has therefore been recognised as the most appropriate and preferred technique to evaluate the predictive accuracy of sexual reoffending, as this method overcomes some of the limits of more traditional methods of predictive accuracy because it is less dependent on and not distorted by variations base rates of recidivism (Mossman, 1994; Rice & Harris, 1995). In general AUC values of .56 correspond to Cohen's *d* value of .2, representing a small effect. An AUC of .65 is equivalent to a Cohen's *d* value of .5, reflecting a moderate effect, whereas an AUC value equal to and above .70 is comparable to Cohen's *d* value of .8, which can be considered a large effect (Rice & Harris, 2005).

Meta-analytic results of the structured professional judgement suggest this approach produces variable results and slightly lower levels of predictive accuracy in comparison to actuarial instruments which achieved considerable support (Hanson & Morton-Bourgon, 2004; Hanson & Morton-Bourgon, 2007). For the prediction of

sexual recidivism, Hanson and Morton-Bourgon (2007) found that the empirical actuarial measures designed for sexual recidivism (*d.* = .70) were more accurate than unstructured professional judgement (*d.* = .43).

Hanson and Morton-Bourgon (2004) found that the predictive accuracies of the actuarial risk scales were in the moderate to large range for predicting sexual recidivism: SORAG (d. = .48), Static -99 (d. = .63), RRASOR (d. = .59), MnSOST-R (d. = .66), and SVR-20 (d. = .77), and that there were no significant differences among the sex offender specific measures with no single risk instrument being significantly more accurate at predicting sexual recidivism.

Appraisal of meta-analyses

Preliminary searches for previous systematic reviews and meta-analyses, identified two meta-analyses (Hanson & Morton-Bourgon, 2004; Hanson & Morton-Bourgon, 2007). The first meta-analysis Hanson and Morton-Bourgon (2004) focused mainly on individual risk factors as predictors of sexual recidivism and then compared actuarial risk instruments with unguided clinical opinion in predicting sexual, violent non-sexual and general recidivism. No significant differences were found in the accuracy of the various actuarial measures. The predictive accuracies for sexual recidivism of the actuarial risk scales were in the moderate to large range, with no significant differences found between specific risk measures (e.g., Static-99, RRASOR or SORAG). This meta-analysis failed to distinguish between predictors for specific subgroups of offenders, for example rapists or child molesters, and focused exclusively on mixed groups of offenders. The second meta-analysis by Hanson and Morton-Bourgon (2007) focused exclusively on the accuracy of different approaches to risk assessment for sexual offender recidivism. The predictive accuracy was found to be moderate to large for actuarial measures.

In Hanson and Morton-Bourgon (2007), the search strategy used in the meta-analysis was based on a limited number of electronic databases, based on four sources; PsyLit, the National Criminal Justice Reference Service (USA), Proquest Digital Dissertations and the library of Public Safety and Emergency Preparedness Canada. However, experts in the field were contacted in order to identify reviews sourced from unpublished studies and data was included from conference presentations. The meta-analysis used wide search terms (child molester, exhibitionism, exhibitionist, failure, frotteur, incest, indecent exposure, paraphilias,

pedophile, pedophilia, predict (ion), rape, rapist, recidivate, recidivism, recidivist, relapse, reoffend, reoffense, sex(ual) offender, sexual assault, sexual deviant) although did not include the term 'risk assessment'. The search was completed in May 2006, since this date further studies have been subsequently been published. The outcome measure for inclusion was any recidivism, including sexual and violent recidivism. Recidivism is a broad definition which can be measured in different ways, and how studies defined and measured recidivism was not differentiated within the meta-analysis. Similarly, different offender types were clumped together under the umbrella term 'sexual offender', when for example, research has shown that some risk instruments are less effective for non-contact offenders and rapists compared to child abusers (Bartosh, Garby, Lewis & Gray, 2003).

Vrieze and Grove (2008) also criticised the findings of Hanson and Morton-Bourgon's (2004) meta-analysis for a faulty statistical approach, arguing that the results cannot be taken at face value because they did not correct for the non-independence of AUC estimates based on common samples. Specific tests need to be combined with the usual independent-samples test they employed, and so this approach may have masked any true differences between mean AUCs and the instruments included within their meta-analysis.

Furthermore the meta-analytical methodology in its ability to address important practical scientific questions remains dependent upon on the strength of the individual studies contained in the analysis. Sources of measurement errors are combined when the variable of interest is defined in different ways in several studies, and then combined into one analysis, making it more difficult to detect a meaningful and significant effect for the variable in question (Lund, 2000). Meta-analysis is not good at identifying multivariant effects and can be biased in the selection of studies included (Kemshall, 2001).

The statistical findings of research studies have been examined within the Hanson and Morton-Bourgon meta-analyses. However, a quality assessment was not undertaken within any of the meta-analysis, attempts were not made to address pertinent aspects of bias in relation to the study design characteristics and an evaluation of the details of the methodological design characteristics and the clarity of reporting within these studies were not examined.

Aims and objectives

To date no systematic review examining the research studies on the predictive validity of risk assessment tools has been conducted. The following review will therefore aim to bridge the gap by evaluating the quality of the literature on the predictive efficacy of risk assessment instruments for men convicted of sexual offences. Utilising a systematic approach to identify studies in this area, this review will summarise and appraise the research findings of actuarial tools, focusing on the structure, sampling and the methodology used for each risk assessment tool. The time frame of the review allowed a rapid approach to be conducted, containing the basic principles of a full systematic review.

Method

Search strategies for identification of studies

An initial scoping exercise was carried out during July 2008 to assess the likely volume of relevant studies, and identified over 7000 references. The author chose to limit the search to references published from 1985 onwards, as actuarial risk assessment tools had not been developed prior to this date. Relevant publications were identified using the following techniques:

a) *Online electronic databases*

OVID PsycInfo - (1985 to wk 2 July 2008, completed on 14 July 2008)
OVID MEDLINE - (1996 to week 1 July 2008, completed on 11 July 2008)
OVID EMBASE - (1988 to week 28 2008, completed on 11 July 2008)
Applied Social Science Index and Abstracts (ASSIA): Cambridge Science Abstracts (CSA) - (1980 to 2008, completed on 17 July 2008)
ISI Web of Knowledge - (1996 to 2008, searched 11 July 2008)

b) *Gateways*

Cochrane CENTRAL (1996 to 2008, completed on July 2008)

c) *Key Meta-analyses and reviews*

Two existing meta-analyses studies (Hanson & Morton-Bourgon, 2004; Hanson & Morton-Bourgon, 2007) were identified as a result of hand-searching.

Search terms

The search terms applied were applied to all databases. The search terms were also modified to meet the requirements of each individual database, depending on differences in fields (see Appendix 1):

> (sex offense* OR sex offence* OR sexual abuse OR Child Molest* OR incest OR rape OR child abuse OR paraphilias OR pedophilia OR paedophilia OR sexual masochism OR sexual sadism OR voyeurism)
> AND
> (risk assessment OR risk management OR statistical validity OR predictive validity OR actuarial OR actuarial analysis OR area under curve OR confidence intervals OR "sensitivity and specificity").

References identified online were imported directly into a database within Refworks. The search syntax can be found in Appendix 1. The author attempted to obtain all publications via the University of Birmingham libraries, interlibrary loans or via direct contact with authors.

Inclusions/exclusion criteria

Inclusion/exclusion criteria were applied to the studies (see Box 1). The inclusion/exclusion form utilised is provided in Appendix 2. If the potential studies adhere to all of the eligibility criteria they will be put forward for inclusion into the final review and quality assessment stage. If the reference is considered to be potentially relevant, a hard copy of the paper will be retrieved for further deliberation.

Box 1: Inclusion/exclusion criteria

Population:	Males sexual offenders over 17 years old.
Intervention:	Actuarial risk assessment tools (designed to specifically to predict risk in adult sexual offenders)
Outcome:	Sexual recidivism, sexual reoffending and sexual reconviction
Study Type:	Cohort and Case Control
Language:	No restrictions
Exclusion:	Narratives, editorials, commentaries and any other type of opinion paper.

This review focused specifically on adult male sexual offender over the age of 17, whom had committed sexual offences. Juvenile sexual offender were not included within this review as risk assessment instruments i.e., the J-SOAP (Prentky, Harris, Frizell & Righthand, 2000) and the ERRASOR (Worling & Cruwen, 2000) have been developed specifically for juvenile populations and not intended for use with adults, the results of which cannot be generalised to other populations.

All actuarial risk instruments identified by the author as being designed for the purpose of predicting future sexual risk were selected. Given the vast array of research in the field of risk prediction and the time constraints of this review, studies will be excluded from the review if the primary focus of the risk assessment tool is not to assess future sexual risk posed or if no evaluation of the instrument has focused on the prediction of sexual risk. Therefore other violent risk instruments such as the PCL-R or VRAG will be excluded, as they were not designed and developed specifically for sexual offenders, despite some studies testing the predictive validity of these instruments of sexual recidivism.

In an attempt to include all relevant publications the outcome measure was set as sexual recidivism, which was broadly defined to include studies using sexual reoffending data measured by new charges, arrests or reconviction. This outcome measure is necessary in order for the predictive validity of risk instruments to be analysed within the studies. This systematic review will not be limited to randomised controlled trials (RCT) as it is anticipated that few if any studies will have been conducted in this area. Case control and cohort studies are typically the type of study

designs used for this type of research. Narrative, editorials and commentaries were also excluded due to not empirically evaluating risk assessment instruments. No language restrictions were imposed, in an attempt to include all relevant research.

The SVR-20 (Boer et al., 1997) was included within this review. Although it is an example of structured professional judgement as opposed to an actuarial tool, some researchers such as Sjöstedt and Långström (2002) and Craig, Beech and Browne (2006) have constructed a score from the checklist based on the transformation of the no, partial and yes categories into 0,1 and 2, so that the SVR-20 could be conceptualised and treated as an actuarial measure.

Assessment of quality

After excluding papers that did not meet the inclusion criteria, the quality of the remaining studies was assessed in two steps:

Step 1- Threshold criteria

The threshold criteria applied were as follows:

- Clear description of risk assessment tools.
- Clear description of outcome measure.
- Sufficient statistical analysis regarding the prediction of sexual recidivism

Studies that did not meet these criteria were excluded for poor quality.

Step 2 – Quality assessment forms

The remaining studies were then assessed using the quality assessment forms. Cohort studies were assessed using quality assessment form A (see Appendix 3) and contained 21 questions. The case control studies were assessed using quality assessment form B (see Appendix 4) which included a total of 27 questions. The questions within the quality assessment forms related to a number of selection bias, measurement/detection bias, attrition bias and other pertinent questions associated with the clarity of the study's results, for example, questions were asked such as whether the risk assessors were blind to the outcome or if inter-rater reliability was reported. The following scoring system was then applied to each of the points; if the conditions were not met then a score of 0 was allocated. If the condition was partially met, a score of 1 was given. When the conditions were fully met a score of 2 was

given. The overall quality score was determined by summing up all scores on each item on the quality assessment form. The higher the score the better quality the study was deemed to be. Finally, if information provided within the study was insufficient to answer the question or failed to be presented clearly enough for a decision to be reached then a separate mark (U=unclear) was awarded. The lucidity of reporting was determined by summing the number of "unclear" items, higher scores indicated less accurate reporting.

A primary reviewer (the author) assessed all 43 studies and a secondary reviewer, a doctoral student, assessed 20% of the studies included to ensure the assessment of quality was consistent.

A Cohen's kappa inter-rater reliability coefficient k = .52 was achieved, and a Cronbach's alpha of .86 (intraclass correlations of .76) was found between the two raters. Fliess (1981) proposed a set of guidelines for interpretation of the Kappa's. Any Kappa's over .75 are considered "excellent", Kappa scores between .6 to .75 are deemed "good" and scores between .4 and .6 are considered "fair/moderate". Given that the inter-rater reliability of this review was found to be within the moderate range, the precise quality score of the studies examined may not be a truly accurate reflection. On the level of discrepancies found ideally a third person should have been consulted, in order to resolve the any disputes which arose. This would have improved the level of inter-rater reliability, unfortunately, due to time constraints this was not possible. It is therefore advisable that some caution is exercised when interpreting the findings of this review, and the conclusions made should be considered tentative.

Data extraction

Data from studies were extracted using a pre-designed data extraction form, detailed in Appendix 5. The quality assessment score and the clarity of reporting score for each individual study was also noted on this form. In cases where information reported was either unclear or sparse, the information was recorded as "not known". Due to the time frame for this review it was not feasible to contact all researchers of those studies.

Results

The full database search yielded a total of 7220 hits (of which 2841 of these references were duplicate publications and two hits were meta-analyses studies). From the total number of hits 7114 studies were then rejected (which included the some of the duplicates and meta-analyses) on the grounds of being irrelevant content to this systematic review.

Of the remaining 106 hits identified, 43 publications were examined and did not meet the inclusion criteria set and so were excluded (see Appendix 6 for reasons for exclusion reasons), leaving a total of 63 potential primary studies. A further 10 publications were subsequently excluded due to the studies being of poor quality and failing to meet the minimum threshold criteria. A further 10 studies could not accessed within the available time frame period and so could not be reviewed. The remaining 43 publications were included for review. Figure 1 shows the search results and the process of study selection.

A number of studies were identified that had not been included within the Hanson and Morton-Bourgon (2007) meta-analysis, which could have been attributable to either differences search terms used or the electronic databases utilised (Barbaree, Seto, Langton & Peacock, 2001; English, Retzlaff & Kleinsasser, 2002; Rice & Harris, 2002; Sjöstedt & Grann, 2002; Craig, Browne, Beech & Stringer, 2004; Seto, Harris, Rice & Barbaree, 2004). A further 8 studies were identified which were published after Hanson and Morton-Bourgon's meta-analysis (Craig, Browne, Beech & Stringer, 2006; Allan, Dawson & Allan, 2006; Allan, Grace, Rutherford & Hudson, 2007; Sreenivasan et al., 2007; Olver, Wong, Nicholaichuk & Gordon, 2007; Langton, Barbaree, Hanson, Harkins & Peacock, 2007; Bengtson & Långström, 2007; Bengtson, 2008).

Description of studies included in the review

Overall the study designs employed by the 43 studies included in the systematic were as follows: 40 cohort studies and 3 case control studies (for a detailed table containing study characteristics see Appendix 7). The total score of AUC values are presented in Table 1.1.

The actuarial measures that were more commonly used with sexual offenders were the Static-99 (Hanson & Thornton, 2000); the Rapid Risk Assessment for Sexual

Offence Recidivism (RRASOR; Hanson 1997); the Sex Offender Risk Appraisal Guide (SORAG; Quinsey et al., 1998); and Minnesota Sex Offender Screening Tool-Revised (MnSOST-R; Epperson et al., 1998). Newer instruments, such as the Static-2002 (Hanson & Thornton, 2003) have received some empirical support but as yet to empirically scrutinised to the same degree as the longer established tools. A total of 10 different risk assessment instruments were included within this review.

Figure 1. Search History

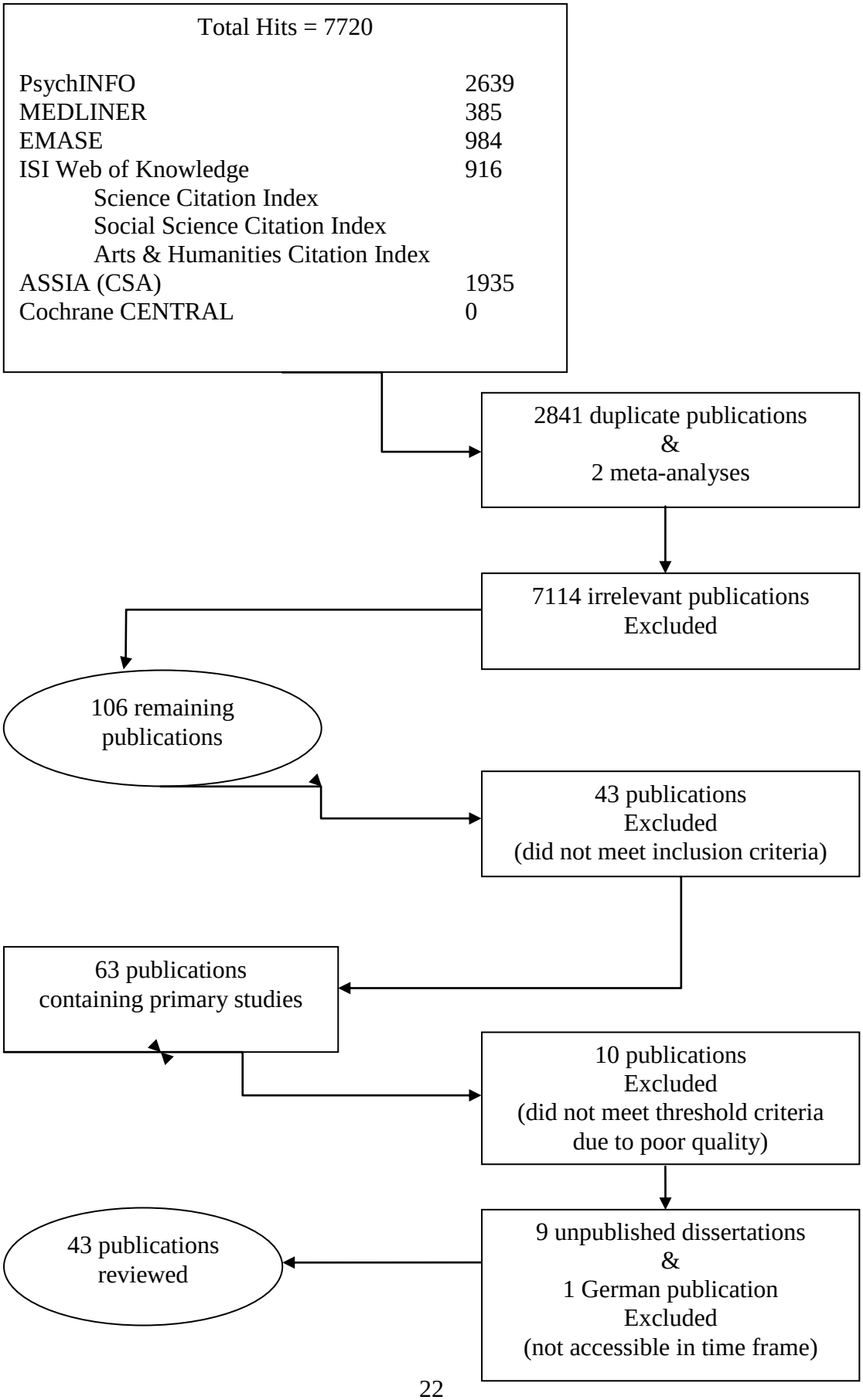


Table 1.1: Summary table showing Area Under the Curve for risk assessment tools

Risk Instrument	Study	Sample Size	Follow-up period	Re-offence rate	Quality Score	AUC[1] scores for sexual recidivism	
RRASOR	Hanson & Thornton (2000)	1301	4-23 yrs	15.4-35.1 %	21	.68	
	Barbaree et al. (2001)	215	4 yrs	9%	30	.77	
	Sjöstedt & Långström (2001)	1400	3.96 yrs	4%	31	.72	
	Sjöstedt & Grann (2002)	1303	6 yrs	3%	24	.73	.94 (imminence) .75 (frequency)
	Sjöstedt & Långström (2002)	51	92.3 mths	20%	25	.73	
	Thornton et al. (2003)	1910	3.1 yrs	2.6%	19	.70	
	Bartosh et al (2003)	186	60-66 mths	11.8%	26	.63 .55 (extra-familial) .73 (intra-familial)	.53 (rapist) .49 (hands-off)
	Harris et al (2003)	396	61.5 mths	26%	34	.59 (total sample) .56 (rapists)	.61 (child molesters)
	Harris & Rice (2003)	No details given	2 yrs and 11 yrs	No details given	18	.71 (IR at 2yr) .70 (no IR at 2yr)	.60 (IR at 11yr) .55 (no IR at 11yr)

Note:[1] AUC value of .50 indicates 'better than chance' prediction, an AUC of >.70 indicates a moderate effect, and an AUC value of >. 75 suggests a moderate to large effect. IR: inter-rater reliability. Imminence: reoffence within one month. Frequency: repeated reoffending. Hands-off: non-contact sexual offenders.

Table 1.1 Continued

Risk Instrument	Study	Sample Size	Follow-up period	Re-offence rate	Quality Score	AUC[1] scores for sexual recidivism	
RRASOR	Seto et al. (2004)	258	5 yrs (study1)	8% (study 1) 19% (study 2)	25	.83 (study 1)	.69 (study 2)
	Långström (2004)	1303	5.7 yrs	No details	34	.73 .76 (Nordic)	.77 (European) .48 (African)
	Allan et al. (2006)	538	9 yrs 3mths	13%	19	.71 .74 (Non-Indigenous)	.65 (Indigenous)
	Langton, Barbaree, Seto et al (2007)	468	5.9 yrs	11%	38	.68	
	Langton, Barbaree, Hanson et al (2007)	464	5.9 yrs	No details	28	.67 (t completers) .71 (t refusers)	.82 (t dropouts)
Static-99	Hanson & Thornton (2000)	1301	4 yrs - 23 yrs	15.4% - 35.1%	21	.71	
	Barbaree et al. (2001)	215	4.5 yrs	9%	30	.70	
	Sjöstedt & Långström (2001)	1400	3.96 yrs	4%	31	.76	
	Beech, Friendship, Erikson & Hanson (2002)	140	6 yrs	15%	22	.77	
	Nunes et al. (2002)	258	7.3 yrs	8.9%	25	.70	
	Thornton (2002)	117	5.8	5.9%	17	.92	
	Sjöstedt & Grann (2002)	1303	6 yrs	3%	24	.75	.75 (frequency) .94 (Imminence)

Note:[1] AUC value of .50 indicates 'better than chance' prediction, an AUC of >.70 indicates a moderate effect, and an AUC value of >. 75 suggests a moderate to large effect. Långström and Allan et al studies compared ethnicity of offenders. t: treatment; IR: inter-rater reliability; Imminence: reoffence within one month; Frequency: repeated reoffending; Study 1: based on 113 child molesters; Study 2: based on 145 child molesters.

Table 1.1 Continued

Risk Instrument	Stud	Sample Size	Follow-up period	Re-offence rate	Quality Score	AUC[1] scores for sexual recidivism	
Static-99	Bartosh et al. (2003)	186	60-66 mths	11.8%	26	.64	
						.65 (extra-familial)	.71 (rapist)
						.74 (intra-familial)	.39 (hands-off)
	Harris & Rice (2003)	No details	2 yrs and 11 yrs	No details	18	.80 (IR at 2 yrs)	.59 (IR at 11 yrs)
						.72 (no IR at 2 yrs)	.59 (no IR at 11 yrs)
	Harris et al. (2003)	396	61.5 mths	26%	34	.62	
						.65 (child molester)	.59 (rapists)
	Craig, Brown, Beech & Stringer (2004)	121	8 yrs	16.2%	27	.59 (at 2yrs)	
						.58 (at 5yrs)	.52 (at 10yrs)
	deVogel et al. (2004)	122	140 mths	39%	29	.71	
	Långström (2004)	1303	5.7 yrs	No details	34	.75	.79 (European)
						.76 (Nordic)	.50 (African)
	Seto et al. (2004)	258	5 yrs (study 1)	8% (1) 19% (2)	25	.81 (study 1)	.72 (study 2)
	Craissati & Beech (2005)	310	4 yrs	2%	26	.71	
						.77 (child molesters)	.53 (rapist)
	Stadtland et al. (2005)	134	9 yrs	9-27.7%	32	.73	
						.74 (non-contact)	.66 (contact)
	Allan et al. (2006)	538	9.3 yrs	13%	19	.71	
						.74 (Non-Indigenous)	.65 (Indigenous)
	Craig, Beech & Browne (2006)	85	8.7 yrs	7%	25	.57 (at 2yrs)	
						.59 (at 5yrs)	.52 (at 10yrs)

Note:[1] AUC value of .50 indicates 'better than chance' prediction, an AUC of >.70 indicates a moderate effect, and an AUC value of >. 75 suggests a moderate to large effect. Långström and Allan et al studies compared ethnicity of offenders. t: treatment; IR: inter-rater reliability. Imminence: reoffence within one month. Frequency: repeated reoffending; Study 1: based on 113 child molesters; Study 2: based on 145 child molesters.

Table 1.1 Continued

Risk Instrument	Study	Sample Size	Follow-up period	Re-offence rate	Quality Score	AUC[1] scores for sexual recidivism	
Static-99	Craig, Browne, Beech & Stringer (2006)	119	72 mths	6% (2yrs)	21	.67 (at 2yrs) .61 (at 5yrs)	.62 (at 10yrs)
	Ducro & Pham (2006)	147	4.2 yrs	25.2%	28	.66 .70 (child abusers)	.71 (rapists)
	Hanson (2006)	3424 (combined samples)	2-23yrs	15.7%	23	.70 (all ages) .68 (age 18-39) .66 (age 40-49)	.76 (age 50-59) .82 (age 60+)
	Looman (2006)	258	5.1 yrs	8.9%	34	.63	
	Stadtland et al. (2006)	134	9yrs	27.6%	n/a	.72 (excluding t dropouts)	.71 (including t dropouts)
	Witte et al. (2006)	72	3yrs	18.3%	20	.72	
	Allen et al. (2007)	495	5.8 yrs	9.9%	24	.72	.81 (Static-99 with deviance)
	Bengtson & Långström (2007)	121	16.3 yrs	31%	37	.62	.72 (severe sexual recidivism)
	Craig, Thornton et al. (2007)	119	6 yrs	6% (2yrs) 12% (5yrs)	20	.66 (at 2yrs)	.60 (at 5yrs)
	Langton, Barbaree, Seto et al. (2007)	468	5.9 yrs	11%	38	.64	
	Langton, Barbaree, Hanson et al. (2007)	464	5.9 yrs	No details	28	.61 (t completers) .82 (t dropouts)	.75 (t refusers)

Note:[1] AUC value of .50 indicates 'better than chance' prediction, an AUC of >.70 indicates a moderate effect, and an AUC value of >. 75 suggests a moderate to large effect. t: treatment; IR: interrater reliability; Severe sexual recidivism: a new sentence for any contact sexual reoffence enforced by physical violence or threats of physical violence. Static-99 with deviance; an overall deviance score was added to the static-99, which was derived from four factors (social inadequacy, sexual interest, hostility and pro-offending attitude) to produce a measure of aggregated dynamic risk.

Table 1.1 Continued

Risk Instrument	Study	Sample Size	Follow-up period	Re-offence rate	Quality Score	AUC[1] scores for sexual recidivism	
Static-99	Knight & Thornton (2007)	599	5 yrs & 10 yrs	21.5%	36	.65	
	Olver et al. (2007)	321	10 yrs	24.6%	35	.63	
	Sreenivasan et al.(2007)	137	13.8 yrs	31%	27	.62	
	Bengtson (2008)	304	16.2 yrs	28 %	30	.64	
						.67 (child molesters)	.64 (rapist)
RM2000/S	Thornton et al. (2003)	1910	3.1 yrs	2.6% (t) 27.7% (ut)	19	.77 (t)	.75(ut)
	Craissati & Beech (2005)	310	No details	2%	26	.70	
						.71 (child molesters)	.67 (rapists)
	Craig, Brown, Beech & Stringer (2004)	122	140 mths	39%	27	.56 (at 2yrs)	
						.58 (at 5yrs)	.55 (at 10yrs)
	Craig, Beech & Browne (2006)	85	8.7 yrs	7%	25	.60 (at 2yrs)	
						.68 (at 5yrs)	.59 (at 10yrs)
	Knight & Thornton (2007)	599	2 , 5 & 10 yrs	21.5% (5 yrs)	36	.63 (at 10yrs)	
	Bengtson (2008)	304	16.2 yrs	28%	30	.65 (total sample)	
						.71 (child molesters)	.61 (rapist)
SVR-20	Sjöstedt & Långström (2002)	51	92.3 mths	20%	25	.49	
	deVogel et al. (2004)	122	140 mths	39%	29	.80	
	Stadtland et al. (2005)	134	9 yrs	27.6 % (contact)	32	.68	
						.54 (non-contact)	.68 (contact)
	Craig, Beech & Browne (2006)	85	8.7 yrs	7%	25	.46 (at 2yrs)	
						.48 (at 5yrs)	.51 (at 10yrs)

Note:[1] AUC value of .50 indicates 'better than chance' prediction, an AUC of >.70 indicates a moderate effect, and an AUC value of >. 75 suggests a moderate to large effect; t: treatment sample; ut: untreated sample.

Table 1.1 Continued

Risk Instrument	Study	Sample Size	Follow-up period	Re-offence rate	Quality Score	AUC[1] scores for sexual recidivism	
SVR-20	Craig, Browne, Beech & Stringer (2006)	85	8.7 yrs	19% (2yrs)	22	.48	
	Dietiker, Dittman & Graf (2007)	64	Not clear	Not clear	n/a	.88	
	Knight & Thornton (2007)	599	2 , 5 & 10 yrs	21.5% (5 yrs)	36	.68	
SORAG	Barbaree et al. (2001)	215	4.5 yrs	9%	30	.70	
	Nunes et al. (2002)	258	7.3 yrs	8.9%	25	.65	
	Rice & Harris (2002)	184	53.6 mths	Not clear	23	.83	
						.81 (extra-familial)	.65 (father-daughter)
	Harris et al. (2003)	396	61.5 mths	26%	34	.66	
						.70 (child molesters)	.62 (rapists)
	Bartosh et al. (2003)	186	60-66 mths	11.8%	26	.58	
						.67 (extra-familial)	.71 (rapist)
						.72 (intra-familial)	.48 (hands-off)
	Seto et al. (2004)	258	5 yrs (study 1)	8% (1) 19% (2)	25	.74 (study 1)	.74 (study 2)
	Ducro & Pham (2006)	147	4.2 yrs	25.2%	28	.64	
						.65 (child abusers)	.64 (rapists)
	Looman (2006)	258	5.1 yrs	8.9%	34	.69	
	Langton, Barbaree, Seto et al. (2007)	464	5.9 yrs	No details	38	.66	
	Knight & Thornton (2007)	599	2 , 5 & 10 yrs	21.5% (5 yrs)	36	.64	

Note:[1] AUC value of .50 indicates 'better than chance' prediction, an AUC of >.70 indicates a moderate effect, and an AUC value of >. 75 suggests a moderate to large effect. Father-daughter: child molesters who molested their daughter/step-daughters. Study 1: based on 113 child molesters; Study 2: based on 145 child molesters.

Table 1.1 Continued

Risk Instrument	Study	Sample Size	Follow-up period	Re-offence rate	Quality Score	AUC[1] scores for sexual recidivism	
MnSOST-R	Barbaree et al. (2001)	215	4.5 yrs	9%	30	.65	
	Bartosh et al. (2003)	186	60-66 mths	11.8%	26	.59	
						.59 (extra-familial)	.54 (rapist)
						.63 (intra-familial)	.57 (hands-off)
	Langton, Barbaree, Seto et al. (2007)	464	5.9 yrs	No details	38	.70	
	Knight & Thornton (2007)	599	2 , 5 & 10 yrs	21.5% (5 yrs)	36	.66	
Static-2002	Bengtson & Långström (2007)	121	16.3 yrs	31%	37	.67	.69 (severe sexual recidivism)
	Langton, Barbaree, Seto et al. (2007)	464	5.9 yrs	No details	38	.71	
	Langton, Barbaree, Hanson et al. (2007)	464	5.9 yrs	No details	28	.69 (t completers)	.84 (t refusers)
						.81 (t dropouts)	
	Knight & Thornton (2007)	599	2 , 5 & 10 yrs	21.5% (5 yrs)	36	.67	
	Bengtson (2008)	304	16.2 yrs	28%	30	.67	
						.69 (child molesters)	.68 (rapist)
SACJ-Min	Hanson & Thornton (2000)	1301	4 - 23yrs	15.4 – 35.1%	21	.67	
CO-SOMB	English et al. (2002)	494	12 – 30 mths	54 % (1 yrs)	13	.64	
VRS-SO	Olver et al. (2007)	321	10 yrs	24.6%	35	.74	

Note:[1] AUC value of .50 indicates 'better than chance' prediction, an AUC of >.70 indicates a moderate effect, and an AUC value of >. 75 suggests a moderate to large effect. t: treatment; Severe sexual recidivism: a new sentence for any contact sexual reoffence enforced by physical violence or threats of physical violence.

Rice and Harris (2005) suggests an AUC value >.70 is indicative of large effect size. Janus and Meehl (1997) argued that actuarial prediction methods need to demonstrate accuracy rates of .70 or .75 as the benchmark for the judiciary to make reliable judgments on civil commitment. On the basis of this, studies within this systematic review will be examined in relation to a recommended threshold of .70.

The Static-99 was the most widely evaluated and used tool in the research literature reviewed, with 32 studies examining the predictive validity, followed by the RRASOR with 13 studies. When all the AUC scores for each individual instrument (based on all offenders within the sample in each study) were compared, only the RRASOR yielded a mean AUC score above the recommended threshold of .70. The SVR-20 had the lowest mean AUC of .64. From Table 1.2 is can been seen that both the RRASOR and the Static-99 had slightly more studies in favour, with overall AUC's above the recommended AUC value of.70, yielding moderate to a large effect and indicating good predictive validity of the risk instruments. The Static-99 produced the widest variability within the results, with a low AUC of .52 (at chance level) and the highest AUC value of .92 reported (see Table 1.2). It can also be seen from Table 1.2 that the RM2000/S, SVR-20, SORAG, MNSOST-R and the Static-2002 had more studies producing AUC values below the recommend threshold than the number of studies supporting the predictive validity of those tools.

Variations were reported in the predictive accuracy when researchers further analysed risk assessment instruments in relation to specific offender characteristics rather than simply looking at the AUC values for the overall sample. It was found for instance that *age* (Hanson, 2006), *ethnicity* (Långström, 2004; Allan et al., 2006), *offender types* (Rice & Harris, 2002; Bartosh et al., 2003; Harris et al., 2003; Craissati & Beech, 2004; Stadtland et al., 2005; Ducro & Pham, 2006; Bengtson, 2008), *treatment status* (Langton, Barbaree, Hanson et al., 2007; Stadtland et al., 2006) and the length of the *follow-up period* (Sjöstedt & Grann, 2002; Craig, Brown, Beech & Stringer, 2004; Craig, Beech & Browne, 2006b) impacted upon the levels of predictive validity.

When researchers divided the samples and examined specific offender types, mixed results were often reported. Risk instruments were generally found to be less accurate for rapists in comparison to child molesters (see Table 1.1). For instance AUC values below .70 were found rapists, but not child molesters (Bengtston, 2008). AUCs also varied for extra-familial and non-contact sexual offenders. Harris, Rice,

Quinsey, Lalumière, Boer and Lang (2003) and Craissati and Beech (2005) have both failed to find support for rapists samples, which was inconsistent with the findings of Bartosh, Garby, Lewis and Gray (2003) and Ducro and Pham (2006), both of whom produced AUC's of .71.

Only two studies, Bartosh et al., (2003) and Stadtland et al., (2005) examined non-contact offenders. Stadtland et al., (2005) found that the SVR-20 was less accurate for non-contact offenders (AUC= .54) in comparison to contact offenders (AUC = .68), whereas the Static-99 was more accurate for non-contact sexual offenders (AUC = .74) compared to contact offenders (AUC= .68).

In terms of ethnicity, only two studies were identified that analysed this variable separately. The RRASOR was less accurate for African offenders and yielded a very low AUC for Australian Indigenous offenders, whereas the Static-99 produced a low AUC for offenders aged between 18-49 years old (see Table 1.1). Only one study (Tough, 2001) using a case control design specifically attempted to validate risk assessment instruments on developmentally disabled sexual offenders. There is a distinct lack of research on the assessment of risk with intellectually disabled sexual offenders and as no other studies attempted to examine intellectual disability, generalisations could not be made. However, preliminary findings from Tough (2001) suggests that the RRASOR correlated with recidivism and was able to distinguish between recidivists and non-recidivists and so could be applied to more diverse populations.

Quality assessment forms were completed on 41 studies of the 43 identified studies, because two of the studies (Dietiker, Dittman & Graf, 2007; Stadtland et al., (2006) could not be quality assessed due to difficulties in translation. Studies were quality assessed on the basis of a number of criteria relating to selection and measurement/detection bias (see Appendix 3 and 4).

Table 1.2: Number of studies showing predictive validity above and below the threshold.

Risk Instrument	Number of studies	Quality score range	Mean quality score	AUC range	Mean AUC[a]	AUC ≥.70	AUC ≤.69
RRASOR	13	18 - 38	27	.59 - .83	.71	8	5
RM2000/S	6	19 - 36	27	.58 - .77	.67	2	4
Static-99	32	17 - 38	27	.52 - .92	.67	17	15
SVR-20	6	22 - 36	28	.48 - .88	.64	1	5
SORAG	10	23 - 38	30	.58 - .83	.68	3	7
MnSOST-R	4	26 - 38	31	.59 - .70	.65	1	3
Static-2002	5	28 - 38	34	.67 - .71	.68	1	3

Note: [a]based on analysis from the AUC of the total treated sample in each study. The mean AUC was calculated based on 40 studies, excluding Tough (2001), Hanson and Harris (2001) and Hanlon, Larson and Zacher (1999) as no AUC's were reported within these case control studies. The SACJ-Min, CO-SOMB and VRS-SO risk instruments were not reported here, given that only one study was conducted on each risk instrument.

Of the studies that were quality assessed, thirty eight of the studies employed a cohort study design and three studies were based on a case-control design. The average quality score for the cohort studies was 26 (SD=6.1) with an average unclear score of 4.1 (SD= 2.4). The quality score ranged from a minimum of 13 to a maximum score of 38 out of a possible score of 40. The number of unclear items ranged from 0 to one study with 8 unclear items. In the case-control studies the average quality score was 30 (SD= 26) with an average unclear score of 4.7 (SD= .6). For the case control studies the scores ranged from 25 to 40 out of a possible score of 52. The number of unclear items ranged from 4 to 5. Studies that were indicative of a poorer quality were those yielding a greater number of unclear items and a lower 'quality' score.

When the AUC values of the risk instruments were compared with the quality scores, the quality scores did not significantly correlate with the AUC's. Only for the Static-99 instrument (r=-0.4, p>.05), it was found that the higher the AUC reported within studies, the lower the quality rating was given.

Table 1.3: Methodological considerations for systematically reviewed studies

	Number of Studies (n= 41[a])		
	Yes	**Partial**	**Unclear/ reported**
Clear definition of sexual recidivism	22 (53.7%)	18 (43.9%)	0
Cohort recruited in acceptable way	18 (47.4%)	14 (36.8%)	4 (10.5%)
Assessor blind to outcome	14 (36.8%)	1 (2.4%)	23 (56.1%)
Assessor trained to complete risk assessments	14 (34.1%)	2 (4.9%)	25 (61%)
Tools coded using multiple sources of information	13 (31.7%)	14 (34.1%)	12 (29.3%)
Inter-rater reliability above .80 threshold	14 (34.1%)	4 (9.8%)	19 (46.3%)
Follow-up period a minimum of 2 yrs[b]	37 (90.2%)	2 (4.9%)	2 (4.9%)
Follow-up period 5 yrs or more[b]	30 (69.8%)	2 (4.7%)	3 (7%)
Missing information recorded and dealt with appropriately	13 (31.7%)	3 (7.3%)	24 (58.5%)
Attrition/participant exclusion recorded	11 (26.8%)	1 (2.4%)	29 (70%)
Large sample size (i.e. over 100)[b]	35 (81.4%)	8 (18.6%)	-
Concurrent validity addressed/discussed	10 (24.4%)	3 (7.3%)	28 (68.3%)

Note: [a] two German studied were not included due to language translation. [b] Including the two German studies as this information was easily extracted.

Of the research reviewed the smallest sample size contained 26 sexual offenders (Hanlon, Larson & Zacher, 1999), whereas the largest sample consisted of 3424 sexual offenders, drawn from 8 different samples (Hanson, 2006). The mean sample size across the studies reviewed was 460. The total number of sex offenders based on the 42 studies was 19,324. Harris & Rice (2003) did not state the number of participants within their study. Some of the participants would, however, have been used in more than one study as researchers have shared and combined samples.

In Table 1.3 it can been seen that on more than half of the studies reviewed it was unclear, or failed to be reported, whether the assessor was blind to recidivism outcome whilst scoring the risk assessment instruments. A further 61% of studies did not clearly record if the assessor(s) were trained in risk assessments. Of the studies reviewed, 32% used more than just file base data (i.e. participant interviews, psychometrics) to complete the risk assessments, however, this is particularly

important when completing risk assessments based on structured clinical guidelines, such as the SVR-20 or SORAG as opposed to tools containing only static risk items.

Over 80% of the studies had large samples of over 100 participants. In terms of long follow-up periods, almost 70% of the studies reviewed were based on five or more years follow-up, only three studies failed to report the follow-up period. Two of the studies (English, Retzlaff & Kleinsasser, 2002; Craig, Thornton, Beech & Browne, 2007) used a mixture of samples, some of which had follow-up periods of five or more years and some sample had shorter follow-ups, and so were given a partial rating score of 1. Using samples with different follow-up periods could potentially skew the data as AUC's have been shown to vary based on length of opportunity to reoffend upon release in the community (Craig, Beech & Browne, 2006b).

Sexual offenders in the studies were selected from a variety of forensic settings, with 42% (18 studies) using sexual offenders from prison samples, 28% (12 studies) using participants from hospital settings or those referred to secure psychiatric hospitals for assessments and 9% (4 studies) were based on community or probation settings, a further 19% (8 studies) used participants from a mixture of settings. In one study (Stadtland, Hollweg, Kleindienst, Dietl, Reich & Nedopil, 2006) the forensic setting was not clear due to difficulties in translation. The mean quality score for the studies based on a prison setting was 26.7 (SD= 6.1), a hospital setting 27.7 (SD= 5.9), a probation setting 28.5 (SD= 7.9) and the mean score for studies using a mixture of different settings was 22.5 (SD= 5.9). However, no significant differences were found between the applied forensic setting and the mean quality review score.

Discussion

This systematic review focused on risk scales rather than individual risk predictors and aimed to investigate whether actuarial risk assessment tools were able to distinguish between recidivists and non-recidivists. The second objective of the review was to assess the quality of the research conducted.

With the exception of newer risk assessment instruments, most of the tools identified within this review have been subject to considerable peer review. The studies reviewed revealed supporting evidence of the validity of the risk tools, with

more studies reported AUC values within the moderate range, consistent with the findings from Hanson and Morton-Bourgon (2007) meta-analysis. Most of the literature on predictive accuracy employed the ROC analysis over other statistical methods, with very few studies reporting odds ratios etc. Research has shown that the predictive accuracy of risk tools was weaker when applied to different cultures (Långström, 2004; Allan, Dawson & Allan, 2006), subtypes of sex offenders (Rice & Harris, 2002; Bartosh et al., 2003; Harris et al., 2003; Stadtland, Hollweg, Kleindienst, Dietl, Reich & Nedopil, 2005) and with older offenders (Hanson, 2006).

Despite endeavours to include all relevant publications using a comprehensive search strategy, this review was subject to a degree of publication bias. Publication bias is the tendency towards research being published when statistically significant results are reported, whilst those studies with negative results fail to get published, thus diminishing the validity of such a systematic review. Attempts were also made to reduce language bias by including papers of all languages, however, two publications in German were identified, unfortunately these papers could not been considered in the quality assessment phase of the review due to difficulties within their interpretation. This systematic review was limited in terms of the time frame within which the review was conducted, and the limited resources precluded the retrieval of unpublished literature or conference presentations.

The quality of the studies reviewed generally reached high standards. However, some systematic biases were found within these studies in terms of selection and measurement bias, which will be discussed.

Selection bias

Selection biases were prominent within the research, with less than half of the studies recruiting samples of all offenders released within a target year or period, reflecting an unbiased cohort. Studies that recruited only offenders who had been referred to assessment and/or treatment at particular institutions contained biases, as there was a lack of information about those offenders not referred, similarly the reasons for such referral failed to be reported.

In relation to the external validity of risk assessment instruments, there have been relatively consistent findings in different westernised countries. Although there has been some replication in continental Europe, most of the research has been conducted in Canada and the United States or the United Kingdom. Of those studies

based on European samples the results have been fairly consistent. From a cultural perspective most risk tools have been designed and validated on North American populations, limiting its applicability for other populations with different cultures. From the limited research that has been conducted using offenders from other cultures such as Australian Indigenous populations (Allan, Dawson & Allan, 2006) or African offenders as seen in the Långström (2004) study, this research suggests that the tools are not as accurate for such populations. The extent to which tools can generalize across samples has previously been questioned because base rates in reoffending are known to vary in subgroups of offenders (Craig, Browne & Stringer, 2004).

Similarly, with regards to offender type there has been a significant lack of empirical evidence to demonstrate the validity of risk assessments with non-contact sex offenders, or those with adult male victims and female sexual offenders. It should be noted that the vast majority of the studies reviewed were also based on non-intellectually disabled sexual offender groups, with only one study reviewed (Tough, 2001) specifically examining sexual offenders with intellectual disabilities, however this study did not use a ROC curve analysis making comparisons with other studies difficult. Whilst studies based on very large numbers of offenders may well incorporate some individuals with intellectual disabilities within their samples, these studies have not controlled for this variable, and so it is difficult to know to the extent how well such risk instruments can be applicable to such specific populations. This highlights the need for further research studies to validate risk instruments on offenders with intellectual disabilities. A further study by Wilcox, Beech, Markall and Blacker (2009) is due to be published on intellectually disabled sexual offenders, which was however not available for inclusion whilst this review was being conducted.

Measurement bias

To examine the predictive validity of risk assessment tools there are a number of statistical techniques, namely the ROC AUC analysis, Pearson's *r* correlation, likelihood ratios, and sensitivity or specificity. Very few studies reported on the sensitivity or specificity, or likelihood ratios. The ROC analysis was the most widely used statistic in the studies reviewed to evaluate predictive validity, however, this statistic is by no means perfect. The ROC analysis cannot compensate for changes in

the risk assessment and cannot remedy problems caused by omitting or changing test items (Harris & Rice, 2003).

Closely linked to the validity of a risk instrument is the reliability of the instrument. High reliability of a risk instrument is necessary for high validity of the instrument. When risk instruments are measured with low reliability, the test cannot be expected to make accurate predictions about risk (Harris & Rice, 2003). Hence, the importance of inter-rater reliability when coding risk assessments, yet most studies failed to report on the reliability of the coding of the individual risk items, or the total scores of actuarial instruments. Furthermore, the ROC method can also not rectify the variability of the duration of the follow-up periods within different studies (Harris & Rice, 2003).

The golden standard for evaluating risk assessment instruments would be to use long-term follow-up periods of convicted sexual offenders who are at large in the community, and without any risk management strategies in place. From an ethical stance withholding supervision from offenders would not be acceptable and so we have to rely on other methodology (Macpherson, 2003). Most of the studies reviewed were based on a retrospective follow-up design when examining sexual risk prediction. This type of study design increased the likelihood of predictor-criterion contamination, especially if raters are not masked to recidivism status. In more than half of the studies reviewed, it was unclear whether the assessor was blind to the outcome measure. The retrospective method does however hold an advantage in that the results can be established very quickly, in contrast to prospective studies whereby it takes many years to collect the data. The retrospective design relies heavily on file based information. Most studies reviewed over-relied on archival file-coding, increasing the likelihood of an underestimation of scores for the measures. Ideally direct evaluation of offenders and the use of multiple sources should be accessed to ensure the quality and reliability of the information. Another issue noted was that of the evaluators training in completing risk assessments. This was either not reported or the level of training/experience not made clear.

With regards to inter-rater reliability, researchers have often failed to systematically include this within their study design. Inter-rater reliability was not completed in 46% of the studies reviewed. When researchers have reported this, the statistical method used to establish reliability was inconsistently used across the studies. Reliability can be reported in three ways, using an average percentage

agreement, intra-class correlation coefficient and Cohen's kappa. When taking percentage agreement it fails correct for chance agreement or account for variance amongst raters (Bohrnstedt & Knoke, 1994). Cohen's Kappa method examines the proportion of raters' agreement whilst correcting for chance agreement but fails to take account of important sources of variance amongst raters (Webster et al., 2006). The intra-class correlation compares the covariance of raters' ratings with the total variance. This was the most widely utilised method in the studies reviewed. A lack of consensus remains as to which is the most effective method in the assessment reliability of the actuarial risk assessment.

The outcome measure used for the inclusion of studies within this review was 'sexual recidivism'. However, there was a lack of consistency and consensus in the operational definition of sexual recidivism incorporated in the studies, which was a further problem for the research in this area. For instance some studies used new reconvictions, some used arrest or charged data, or the repeated return to a forensic unit. In Tough's (2001) study sexual recidivism was defined any sexual or nonsexual charge or conviction and any sexual misbehaviour reported by other 'official' sources such as agency staff or other professionals, or verbal complaints from community members. More than half of the studies reviewed used very broad definitions and did not specify the types of offences, or incorporated all sexual offences into the definition, and some specifically excluded any non-contact sexual offences, such as exhibitionism or child pornography in their definitions of sexual recidivism. Thirteen studies (32%) including both contact and non-contact sexual offences, however, twenty (49%) of the studies reviewed were not specific in the offence inclusion criteria for sexual recidivism. Studies were therefore not always measuring the same construct, making meaningful comparisons between studies difficult. Such variations in definitions may account for differences reported in the AUCs between studies.

Further problems are inherent when studies rely on 'reconviction' as an outcome measure as this is an underestimation as many sexual offences go unreported. Serious sexual offences can often be bargained down to violent offences in order to secure conviction (Bagley & Pritchard, 2000). Similarly resistance within the legal system to investigate sexual offence incidents when committed by individuals with intellectual disabilities has resulted in a lesser likelihood of legal prosecution (Swanson & Garwick, 1990). With this in mind, it is advisable for researchers to incorporate a broader definition of sexual recidivism, and not simply

rely on reconvictions. For instance, within the UK the use of official sources for reconviction data has been criticised, as a poor correlation was found between the Offender Index and the National Identification Service for offence based criminal history summaries (Friendship, Thornton, Erikson & Beech, 2001). Most researchers only cited one source for collected recidivism data. Reconviction data as a dichotomous variable (i.e. reconvicted/not reconvicted) fails to account for the frequency or severity of re-offending (Friendship, Beech & Browne, 2002; Beech, Fisher & Thornton, 2003). Litwack (2001) recommends the use of clinically relevant outcome criteria such as imminence, frequency and severity of reoffending. This was subsequently examined by Sjöstedt and Grann (2002), who found a higher predictive accuracy when using imminence (reoffence within one month of release) as an outcome measure.

Other notable factors that could account for variations in the predictive accuracy of risk instruments could be attributable to the variability of follow-up periods and the amount of missing data. The average length of opportunity to reoffend does affect the AUCs. Within the studies reviewed the shortest mean length of follow-up examined was two years and the longest follow-up period recorded was 23 years (Hanson & Thornton, 2000). Such variations make comparison difficult, although it should be noted that the majority of the studies reviewed upheld appropriate follow-up periods of five or more years.

Whilst it could be assumed that risk assessments were conducted blind to the recidivism status, this was occasionally not stated, reflecting a lacked clarity in the reporting of methodology by some researchers in this field.

Recommendations for future research

More replication of the research is needed using settings and populations from other non-western countries, and using different offender populations, such as sexual offenders with intellectual/learning disabilities. Research would benefit from separating offender types instead of analysing all offenders together, such as those with only non-contact offences or adult male victims.

Future research would also benefit from the use of more than one outcome measure, using both a broad and consistent definition of sexual recidivism to counter the problems highlighted above. The inclusion of inter-rater reliability, particularly for

tools such as the SVR-20 is paramount, as well as the consistent use of the same method of inter-rater reliability analysis, is also required in future research.

Conclusions and implications for practice

On the basis of the studies reviewed the predictive validity of the risk assessments for general sexual offenders fell within the moderate range. When age, ethnicity and the subtype of sexual offender were accounted for, this was found to influence the effectiveness of the tools in the ability to predict future sexual reoffending. Risk tools were better able to predict future risk of reoffending in child molesters than rapists, older offenders than younger offenders and were generally less effective when applied to ethnic minorities.

Researchers were neglectful when it came to examining the reliability of risk assessment tools, with a lack of inter-rater reliability being reported the reliability of the findings could be called into question. Clarity within the methodological reporting is therefore paramount.

Professionals assessing the risk posed by sexual offenders need to be aware of the limitations of current risk instruments. As the number of empirical studies assessing risk assessment has grown in recent years, clinicians' confidence and understanding of the strengths and limitations of the risk assessment instruments they apply has also increased. Whilst more is known regarding the efficiency of tools with general sexual offenders, conclusions regarding the generalisability of the findings to other diverse sexual offender populations remain uncertain. Caution should still remain when applying actuarial tools for populations of different ethnic minorities, or when examining specific populations such as offenders with intellectual disabilities, given the lack of knowledge regarding these groups at this stage.

Chapter Three

The Assessment of Static and Dynamic Risk and Recidivism in a Sample of Special Needs Sexual Offenders

The previous chapter highlighted the importance of accurate risk assessment tools and noted the requirement for additional high quality studies to validate risk assessment instruments on subgroups of offenders, in particular sexual offenders with intellectual disabilities. The majority of research on risk assessments to date has focused exclusively on mainstream samples and there is a lack of a comprehensive review of the current state of risk assessments with offenders for whom mainstream assessment methods may not be appropriate. Research has provided support for the predictive validity of static actuarial risk assessment tools (Hanson & Morton-Bourgon, 2007), however, dynamic risk factors have been brought into the arena and have been the focus of attention in research in recent years, due to the limitations of static instruments outlined in the introduction to this thesis.

Appropriate risk assessments need to be made available for an intellectually disabled population, a problem this current chapter aims to address, by investigating the predictive validity of three current risk assessment instruments used with mainstream sexual offenders on a sample of offenders with intellectual disabilities. This is also the first study to examine the accuracy and utility of the Assessment of Risk Manageability for Intellectually Disabled Individuals who Offend (ARMIDILO; Boer, Tough & Haaven, 2004), a new dynamic risk assessment instrument specifically designed for the needs of intellectually disabled sexual offenders.

Abstract

Background The predictive validity of four risk assessment instruments: the RRASOR, SVR-20, RM2000-V and the ARMIDILO Stable, and Acute dynamic client subscales were assessed on a sample of 88 offenders: 44 mainstream and 44 sexual offenders with special needs, who had been matched on risk items within the RRASOR tool.

Method Instruments were coded retrospectively from file information. Sexual reconviction data was used, in conjunction with sexual recidivism data based on unofficial data sources, over a mean follow-up period of 8.8 years.

Results The results of this study found that the ARMIDILO instrument was the best predictor for sexual reconviction among offenders with special needs (ARMIDILO-Stable, AUC = .60; ARMIDILO-Acute, AUC = .73). While, the predictive validities of the RRASOR (AUC = .53) and the RM2000-V (AUC = .50) were little better than chance. In contrast, the SVR-20 yielded a higher score (AUC =.73) for the non-ID sample, than for the intellectually disabled sample (AUC =.45). Within the special needs group, the ARMIDILO-Acute, SVR-20 Psychosocial Affect, and Overall scales were better predictors of sexual recidivism for the intellectually disabled subgroup (AUCs ranging from .75 to .88).

Conclusions Static risk assessment instruments were more accurate predictors of sexual recidivism for mainstream sexual offenders than for offenders with special needs. However, the ARMIDILO instrument was better able to predict future risks in offenders with special needs, suggesting that dynamic risk factors are pivotal in the assessment of risk for this population. These preliminary results of the ARMIDILO tool are promising but further empirical validation is still required.

Introduction

Researchers have invested a great deal of effort to determine the effectiveness of risk assessment tools, validating them against large samples of mainstream sexual offenders. Unfortunately, sex offenders with intellectual disabilities have been largely overlooked, and very few studies to date have examined the predictive validity of different risk assessment tools, or approaches. This leads to the question of whether such tools should be applied to this specific population of sexual offenders.

Research on predicting reoffending behaviour of those with intellectual disabilities has been limited in comparison with the quantity of mainstream research (Harris & Tough, 2004; Lambrick, 2003; Lindsay & Beail, 2004). In the prediction of violent recidivism Grey et al. (2007) found that the VRAG, PCL-SV and the HCR-20 were all significant predictors in violent reconviction in a sample of offenders with intellectual disabilities, suggesting that such risk assessment tools are efficacious with such a population. In relation to the assessment of sexual risk a consensus has yet to be established as to how applicable risk assessment tools are for offenders with intellectual impairments. There remains a pressing need for future research into risk assessment of intellectually disabled sexual offenders (Craig & Hutchinson, 2005). Currently, the established formal measures of risk for mainstream populations do not take into account specific issues relevant to the intellectually disabled population, and no reliable static actuarial measures have been developed, which have been normed specifically on a population of sex offenders with intellectual disabilities (Lambrick & Glaser, 2004). Without reliable valid instruments, clinicians may develop their own idiosyncratic risk assessments that have not been empirically validated in terms of predictive validity. This makes comparisons between individuals meaningless, as one service may judge a client to be a high risk and another service may perceive the same person to be a moderate risk (Lindsay & Beail, 2004).

It remains to be determined whether it is appropriate for current actuarial measures to be administered to those with intellectual disabilities. Craig and Hutchinson (2005) have argued that the extent to which risk measures can be applied to diverse populations is currently unclear. This is because the actuarial method requires a comparison of the similarities of an offender's profile to aggregated knowledge of past events. In relation to this, Craig and Hutchinson (2005) suggest that a reduction in the predictive accuracy of the scale may occur when applied to an

individual with characteristics that differ from the data cohort, notably with regards to base rate.

The legal system is often reticent about, and resistant to investigating sexual offence incidents committed by individuals with intellectual disabilities, thus reducing the likelihood of legal prosecution, which subsequently affects base rate data (Swanson & Garwick, 1990). Furthermore, sexual offenders with intellectual disabilities are more likely to be dealt in the Mental Health System than by the Criminal Justice System, depending on the identity of the victim (i.e. a prosecution is more likely to occur if the victim is male and a child) than on the nature of the offence (Green, Grey & Willner, 2002).

Harris and Tough (2004) argue that while there is an absence of alternative risk assessment instruments directly for this population, there is no research that vindicates that there are risk predictors that are any different for those with intellectual disabilities compared with general offenders, and that it is therefore appropriate and advisable to use actuarial assessments that have been standardised on the general sex offender population. Similarly, Wilcox (2004) argues that risk factors contained within risk assessment tools can apply to intellectually disabled sex offenders, given that initial research to identify risk factors associated with reconviction was based on large offender populations, which would have been normally distributed in terms of intelligence, and that subsequently, a proportion of them would have had intellectual disabilities.

On inspection of the items contained within some of the widely used actuarial tools, the Risk Matrix 2000 (RM2000, Thornton et al., 2003), for example contains risk factors such as having a male victim or relationship status. The presence of such factors incrementally increase the risk category, however research has shown that the offence characteristics of men with intellectual disabilities are different and that they tend to offend against more male victims than non-disabled sex offenders, and their offences tend to be less serious (Brown & Stein, 1997). Similarly, sexual offenders with intellectual disabilities have been found to have poor peer relations, a lack of social sexual knowledge, negative early sexual experiences, and a confused self-concept (Hayes, 1991).

Day (1994) examined the profile of sexual offenders with intellectual disabilities and found that prominent features of this group of offenders included sexual naivety, poor impulse control, inability to understand normal sexual

relationships, and a lack of relationship skills. Circumstance and opportunity rather than sexual preference appeared to be the overriding factors in the choice of victims in the majority of cases. Lacking an emotional relationship would therefore instantly place intellectually disabled offenders in a higher risk category. If offenders with intellectual disabilities are indeed characteristically different and represent a homogeneous group, then how does this impact on the assessment of risk?

Research conducted to examine the characteristics of intellectually disabled sex offenders has, however generally been methodologically flawed, on the basis of using only clinical trials with no control group in which to show that these characteristics do not exist in other samples of individuals with an intellectual disability (Lindsay, 2002).

Harris and Tough (2004) argue that the Rapid Risk Assessment of Sexual Offence Recidivism (RRASOR; Hanson, 1997) instrument has practical utility, and suggest that it is a useful metric of risk for sex offenders with intellectual disabilities, based on unpublished research by Tough (2001). The RRASOR was found to perform better than the most well known actuarial measure Static-99 (Hanson & Thornton, 2000), and was able to differentiate recidivists from non-recidivists, having good estimates for overall risk based on a sample of 76 intellectually disabled sex offenders. Tough (2001) therefore argues that actuarial risk assessment can be used to enhance supervision and rationalise treatment. However, Tough's statistical analysis was limited to t-tests comparing the recidivists with non-recidivists. A Receiver Operating Characteristic (ROC) Area Under the Curve analysis (AUC) was not used to evaluate the overall predictive efficacy of these risk assessment tools. The AUC analysis measures discrimination, i.e. the ability of the risk assessment to correctly identify those who will reoffend in the future. The ROC curve is the preferred index used to evaluate the predictive accuracy of a risk assessment instrument and has the advantage in that it is insensitive to variations in base rates (Mossman, 1994).

More recently, Lindsay et al. (2008) investigated the accuracy of the RM2000-C (combined risk) scale and the Static-99 on 212 offenders with intellectual disabilities. Static-99 was shown to have a significant AUC of .71, whereas the RM2000-S fell short of significance and produced an AUC of.61. This study did not use official recorded offences as an outcome measure and recidivism data was collected in a short time span of one year. However, Wilcox, Beech, Markall and Blacker (2009) compared three mainstream risk tools: the RRASOR, Static-99 and

RM2000-Sexual on a sample of 27 treated intellectually disabled sex offenders, and found that the Static-99 had a lower AUC than found in a study by Lindsay et al. Overall, the Static-99 had the highest AUC of .64, followed by the RM2000-S (AUC of .58), which was lower than that found by Lindsay et al. The predictive validity of the RRASOR produced the poorest score (AUC of .42).

At present it would appear that the Static-99 is the most accurate of these tools and may be used with sexual offenders with intellectual disabilities. These results are inconsistent with Tough (2001), which may be attributable to differences in statistical procedures and differences in the severity of the intellectual deficits, as the participants in Tough's sample all had significant cognitive deficits with a diagnosis of 'mental retardation' whereas the Wilcox et al., sample included those with mild intellectual disabilities.

For a more comprehensive risk assessment, Harris and Tough (2004) advocate the use of dynamic risk assessments in combination with static actuarial measures, as the static factors would yield a baseline appraisal and the dynamic factors would monitor changes in the level of risk over time and direct treatment intervention. Similarly, Craig et al. (2005) suggest using actuarial estimates in conjunction with dynamic change and psychometric measures.

Previously, most of the research relating to dynamic risk factors has been based on mainstream populations. However, a study conducted by Lindsay, Elliot and Astell (2004) has provided confirmation that dynamic factors are of value, and are equally as applicable to the sexual offenders with intellectual disabilities. In a retrospective comparison study, Lindsay et al. (2004) selected a range of predictive variables from the literature that had had empirical support. They found that antisocial attitudes, denial of crime, erratic attendance, and poor response to treatment were dynamic factors that significantly predicted recidivism. Interestingly other dynamic variables such as criminal lifestyle, criminal companions, social and emotional isolation and mental illness were not found to be related to recidivism.

Replication of the results are vital in determining which dynamic risk factors are most relevant for intellectually disabled sex offenders. However, it should be noted that research into dynamic risk predictors is still in its infancy, and there remains a lack of consensus based on mainstream offenders, as to which dynamic risk factors account for the most variance in predicting sexual recidivism (Craig et al., 2005).

Dynamic risk factors have been included in checklists based on the structured professional judgement approach to risk assessment. However, no research has examined the predictive validity on an intellectually disabled population. Boer, Tough and Haaven (2004) presented a convergent approach, whereby they suggest the inclusion of both information that is gained by static actuarial instruments, and of chronic and acute dynamic factors, in the assessment of risk management with intellectually disabled sex offenders. Boer, Tough and Haaven (2004) developed the ARMIDILO (Assessment of Risk Manageability for Intellectually Disabled Individuals who Offend) checklist containing 30 stable and acute dynamic risk factors. However, the ARMIDILO has yet to be empirically validated.

In addition to dynamic and static factors in the assessment of risk, Haaven (2005) recommended that environmental factors need to be added to the assessment of risk for this particular population of sexual offender. With appropriate supervision, based on a good understanding of the issues and the ability to intercept at proper times, Haaven argues that actuarial assessment would give a much higher level of risk in the individual than they actually are.

Research aims

The predictive value of the SVR-20 (Boer, Hart, Kropp & Webster, 1997), RM2000-V (Thornton et al., 2003), RRASOR (Hanson, 1997) and ARMIDILO (Boer, Tough & Haaven, 2004) instruments was examined using a sample of intellectual disabled sex offenders, and matched with a sample of non-intellectually disabled offenders. It was hypothesised that:

1. Risk assessment tools would assist in estimating the level of risk when applied to a sample of intellectual disabled sex offenders.
2. The SVR-20 and the ARMIDILO would perform better than the RRASOR (given that they incorporate a number of potential risk variables that are good at capturing problems in intellectually disabled offenders).

Method

Participants

The sample consists of 88 sexual offenders (44 special needs sexual offenders and 44 matched mainstream/non-ID sexual offenders). The non-ID offenders were selected from a list of offenders who participated in treatment during the same period as the special needs group. The non-ID offenders were only then included if they provided a corresponding match based on every item on the RRASOR, this was to ensure that the mainstream and intellectually disabled group were rated the same risk category at the time of completing treatment (See Table 2.1). If offenders are considered the same risk level on static actuarial instruments and the instruments are considered to be effective, then the predictive validity between the groups should be comparable. On the RRASOR three (6.8%) offenders obtained a score of 0, 14 (31.8%) offenders obtained a score of 1, 15 (34.1%) were given a score of 2, six (13.6%) were given a score of 3 and another six (13.6%) were given a score of 4. No offenders reached a score of 5 on the RRASOR.

Both the special needs and mainstream offenders participated in a treatment group between 1994 and 2001, at the West Midlands Probation Service as a condition of their Probation Order. The length of follow up and the static risk were therefore equivalent in both groups (See Table 2.1). The average follow-up period was the amount of time between the first opportunity to reoffend and the last follow-up information, which was collected in June 2008. The special needs sample was derived from an adapted sex offender treatment programme, for offenders with intellectual disabilities to borderline levels of intellectual functioning (IQ 70-80). The treatment programme was designed for those offenders deemed not suitable for inclusion in the mainstream programme due to intellectual difficulties and significant literacy deficits. This sample therefore represented a diverse group rather than a strictly intellectually disabled population.

It should be noted that the mean IQ score is not based on the total sample, only those who received a standardised test (n =14). A very small number of the sample (n = 4) had an IQ in the borderline level of functioning range. It can be seen from Table 2.1 that the offence characteristics of the non-ID group were comparable to the special needs groups.

Table 2.1. Sample Characteristics

	Intellectually Disabled Group			Non-ID Group		
	%	*No.*		*%*	*No.*	
Gender of Victim						
Female	75%	33		72.7%	32	
Male	21.5%	9		22.7%	10	
Both	4.5%	2		4.5%	2	
Type of Victim						
Adult	20.5%	9		27.3%	12	
Child	70.5%	31		68.2%	30	
Both	9.1%	4		4.5%	2	
Unrelated to Victim(s)	93.2%	41		93.2%	41	
RM2000V Score						
Low(0-1)	11	25		22	50	
Med (2-3)	26	59.1		18	40.9	
High (4-5)	6	13.7		3	6.8	
Very High (6+)	1	2.3		1	2.3	
	Range	*Mean*	*SD*	*Range*	*Mean*	*SD*
Treatment Length (months)	30 (3-32)	18	8.8	56 (1-57)	17.7	11.2
Follow-up Period (months)[a]	93 (54-147)	109	23.0	75 (68-143)	101	15.3
Age[b] (years)	63 (19-82)	35	14.80	45 (19-64)	38.5	12.4
IQ score	21 (58-79)	69	8.3	27 (92-119)	107	7.5

Note. [a]The average follow-up period was the amount of time between the first opportunity to reoffend and the last follow-up information , collected in June 2008. [b] Age was recorded from the age at completion of treatment. Risk scores were only provided for the RM2000-V, as both the RRASOR, SVR-20 and the ARMIDILO instruments do not have risk category levels. A simple summation of the items scores would not be appropriate as a final risk judgement needs to be derived from integrating all available information, as the presence of one or two items could be sufficient to justify a "high risk" judgement. Risk categories have therefore not been provided for these risk instruments.

Assessment of intellectual functioning

Not all of the offenders received the same standardised test. Initially, intellectual functioning was established using clinical judgement, with a later introduction of formal assessments. The Shipley Institute of Living Scale (SILS) was the main measure of IQ. The SILS is designed to assess general intellectual functioning in adults and adolescents and to aid in detecting cognitive impairment in individuals with normal intelligence. The scale consists of two sub-tests, a 40-item vocabulary test and a 20-item test of abstract thinking and measures the discrepancy between vocabulary and abstract concept formation. In some cases the Wechsler Adult Intelligence Scale – Revised (WAIS-R) was also used. The SILS provides an estimated equivalent Full Scale WAIS score, using conversion tables supplied in the test manual (Zachary, 1986).

Procedure

The current study involved retrospective file coding of archival files. The risk instruments were coded on the basis of information contained within the files, such as treatment progress, treatment plans or psychological reports and psychometric assessments. The risk assessments were completed around the time that the offenders had completed treatment. In cases where file information was limited, a consultation with treatment facilitators was held to gain the relevant information regarding dynamic factors. To minimise the problem of missing information, two treatment facilitators were interviewed. The probation files in this study contained information ranging from psychological reports, psychiatric reports, legal reports (such as the Pre-Sentence Reports, victim/witness statements and police reports), as well as treatment notes and psychometric tests. In some instances, Polygraph Assessments were also contained within the file.

Measures

Sexual Violence Risk Scale (SVR-20; Boer et al., 1997)

The SVR-20 adopts a guided clinical assessment format to assess the risk of sexual violence recidivism. The SVR-20 contains 20 items of empirically validated risk factors, which are divided into three risk categories: 1) psychosocial adjustment; 2) sexual offences; 3) future planning. Stable and acute dynamic factors are included in the SVR-20. Items are coded on a 3-point scale based on the method described by

Sjöstedt and Långström (2001), where "0" indicates that the item does not apply according to the available information; "1" the item probably or partially applies, and "2" indicates that the item definitely applies. The final risk judgement is indicated by the following bands: Low (score 0-1), Moderate, (score 2-5), High (score 6 or more). The SVR-20 had not been validated on an intellectually disabled sexual offender population and was also available to West Midlands Probation Services at the time research commenced. The RSVP (Hart, Kropp, Laws, Klaver, Logan & Watt, 2003) whilst more recent, is not a revision of the SVR-20 but does cover elements missed in the SVR-20, however, it is not geared towards risk prediction, but rather a system of managing the immediate risk and providing guidance for risk formulation (Craig, Browne & Beech, 2008). The RSVP was also considered more demanding in the time taken to complete the tool and so the SVR-20 was selected over the RSVP.

Risk Matrix 2000 (Thornton et al., 2003)

The RM2000 is a two-dimensional risk assessment for sex offenders. The first scale is the Risk Matrix 2000/Sexual (RMS), measuring the risk of sexual recidivism, and the second scale is the Risk Matrix 2000/Violent (RMV) measuring the risk of non-sexual violence. The RM2000 scales use a step-wise approach to scoring. The RM2000/V scale contains only three items which include age upon release, the amount of prior violence and a history of burglary. The RMV was found to significantly predict violent recidivism in sexual and sexual/violent offender groups, with AUCs of .85 and .85 respectively (Craig, Beech & Browne, 2006b). More recently in the study by Wilcox, Beech, Markall and Blacker (2009) the RMS scale had been examined on a sample of intellectually disabled sexual offenders and found to be little better than chance, with an AUC of .58. The RMV however, has not been validated on an intellectually disabled sample and so was selected as a measure for assessing violent risk.

Rapid Risk Assessment for Sexual Offence Recidivism (RRASOR, Hanson, 1997)

The Rapid Risk Assessment for Sexual Offence Recidivism (RRASOR) is an actuarial tool containing four items. These items are past sexual offences, age at commencement of risk, extrafamilial victims, and male victims. Offenders are allocated points according to the presence of these, and are given a risk categorisation

on this basis. The RRASOR had received extensive empirical support for distinguishing between sexual recidivists and non-recidivists (Barbaree et al., 2001; Bartosh et al., 2003; Hanson & Thornton, 2000; Harris et al., 2003; Sjöstedt & Långström, 2001; Långström, 2004).

Assessment of Risk Manageability for Intellectually Disabled Individuals who Offend (ARMIDILO; Boer, Tough & Haaven, 2004)

The ARMIDILO incorporates 30 items including *stable* dynamic factors, the relatively persistent characteristics of the offender; and *acute* dynamic factors, the rapidly changing contextual factors that signal the onset of offending. *Stable* and *acute* dynamic factors were further divided into environmental and client related factors: 1) *stable dynamic environmental* subscale (attitudes towards intellectually disabled individuals, communication amongst staff, client specific knowledge, consistency of supervision, situational consistency and generalisation of treatment concepts); 2) *acute dynamic environmental* subscale items (changes in social relationships, new supervisory staff, monitoring of offender by staff, victim access, situational changes and access to intoxicants); 3) *stable dynamic client subscale* (attitude towards and compliance with treatment and supervision, knowledge of faulty thought, behaviour cycle, risk factors and relapse prevention plan, sexual knowledge and self-management, mental health problems, time management skills, substance abuse, victim selection, general coping ability and self-efficacy, use of violence towards self or others, and finally impulsivity); 4) *acute dynamic client* subscale (changes in substance abuse pattern, sexual preoccupation, emotional state, victim access, attitude or behaviour in treatment or supervision and changes in the ability to use coping strategies). Items were scored on a 5-point scale ranging from -2 reducing risk to +2 indicating an increase in risk. According to the authors, the ARMIDILO is still in its construction phase and lacks any empirical validation. Only the *acute* and *stable dynamic client* subscales were used within this study as information required to complete the environmental subscales was not documented within file information.

Outcome measures used in the study

As sexual offenders with intellectual disabilities often escape prosecution for illegal sexual acts (Swanson & Garwick, 1990), it was considered appropriate to include the

broader definition of sexual recidivism rather than reconviction data alone, by examining both unofficial sources of data in conjunction with official rates of sexual reconviction. To ensure quality and reliability of the results, two outcome measures were therefore included in this study: the conviction of a new sexual offence, and sexual recidivism behaviour.

Sexual reconviction was defined as a subsequent conviction for any sexual offence that occurred after the completion of a sex offender treatment programme. Sexual recidivism in this study was based on the definition provided by Falshaw, Friendship and Bates (2003), where sexual recidivism was referred to as any offence-related behaviour, either legal or illegal, with a clear sexual motivation, e.g., a convicted child sexual offender found loitering outside a primary school or in a park.

The outcome measure was sexual reconviction rates, collected from the Home Office Police National Computer (HOPNC) database, in order to determine a group of recidivists and a group of non-recidivists. The unofficial sexual recidivism data was gathered from any information contained within probation files, which included any cautions or warnings for sexual reoffending, or any recorded information relating to inappropriate sexual or grooming behaviours, which may have been documented in memos from telephone messages or treatment notes. It should be noted that the sexual recidivism data was only collected during the time span whilst the offenders were on Probation Orders, and not collected after the completion of treatment. As the unofficial sexual recidivism data was sparse it was not used in isolation but was combined with the official reconviction data from the HOPNC.

Ethics Procedure

Ethical approval was obtained from West Midlands Probation Services in January 2007. The ethics form was completed detailing the aims of the study. The participants remained anonymous and provided a code number when confidential data was entered into SPSS. Given the confidential and sensitive nature of the information, all information was kept in a locked filing cabinet within a locked room at the University of Birmingham. The data will be kept for five years and then destroyed.

Data Analysis

The predictive validity of the risk scales was derived using the Relative Operating Characteristics (ROC) Area Under the Curve (AUC) analysis. This is the most

commonly used technique in the field of risk prediction (Mossman, 1994) and allows the evaluation of the sensitivity-specificity (i.e., true positive vs. false positives) trade-off at all possible cut off scores, independent of the variations in base rates. The AUC can be interpreted as the probability of a randomly selected recidivist scoring higher on a risk instrument than a randomly selected non-recidivist. An AUC of 0.5 indicates prediction at chance level and an AUC of 1.0 indicates a perfect discrimination. Statistical analysis was performed using SPSS 15.0. Associations between risk scores and sexual recidivism and reconviction were analysed with Pearson's (*r*) coefficients and t-tests were conducted to compare difference between the special needs group and the mainstream comparison group on parametric data for the tests.

Results

It was found that the special needs group were reconvicted at a similar rate to the mainstream group (special needs: 25%, n = 11; mainstream 30%, n =14). There was also no significant difference between the unofficial recidivism rate of the two samples (special needs: 46%, n = 20; mainstream 32%, n =14). Similarly, no significant differences were found in violent reconviction rates (special needs: 11%, n = 5; mainstream 5%, n =2); and non-sexual/non-violent reconviction, such as theft, driving offences etc. (special needs: 31%, n = 14; mainstream 27%, n =12).

Comparisons of the offenders with special needs and mainstream offenders on the risk assessment instruments employed in the study are outlined in Table 2.2. Statistical comparisons of the groups revealed that the special needs group had significantly higher ARMIDILO Stable scores, $t(88) = 3.12$, $p < .05$, in comparison to the non-ID group. The special needs group also had significantly greater scores on the SVR-20 Psychosocial scale, $t(86) = 2.71$, $p < .05$ and on the SVR-20 Overall scale, $t(86) = 2.86$, $p < .05$, in comparison to the non-ID group. However, the special needs group did not differ significantly on the RRASOR, SVR-20 Future plans or the ARMIDILO Acute scales.

Comparisons between the recidivist and the non-recidivist samples (across both the special needs and non-ID groups) found no difference in the overall scores on the SVR-20 and the RRASOR. The sexual recidivists had a higher mean score on the SVR-20 Overall scale ($M = 15.08$, $SD = 4.80$) compared to the non-recidivist group ($M = 13.44$, $SD = 7.20$), but this difference was not significant. Similarly the recidivist

sample had a higher mean score (*M* = 2.21, *SD* = 1.17) on the RRASOR than the non-recidivists (*M* = 1.86, *SD* = 1.11), again this difference was not significant.

Predictive validity analyses

The AUCs for the risk scales are displayed in Table 2.3 and 2.4. The AUCs for unofficial sexual recidivism, sexual, violent and general reconvictions are shown in Table 2.3 for the special needs and non-ID groups. Table 2.4 presents the AUCs for sexual reconviction for the special needs group only.

Table 2.2: Descriptive Statistics of the RRASOR, RM2000-V, SVR-20, and the ARMIDILLO scales for the ID and Non-ID Groups

Risk assessment instrument	ID subgroup						Special Needs group (n=44)			Non-ID group (n=44)		
	ID (IQ <75) (n=10)			ID Borderline (IQ >75) (n=4)								
	M	SD	Range	M	SD	Range	M	SD	Range	M	SD	Range
RRASOR	2.10	1.20	4 (0 - 4)	2.00	.82	2 (1-3)	2.01	1.14	4 (0-4)	1.95	1.14	4.0 (0-4)
RM2000-V	2.20	1.03	4 (0 - 4)	2.00	1.41	3 (0-3)	2.25	1.37	6 (0-6)	1.66	1.60	6 (0 -6)
SVR-20 Psychosocial Adjustment	9.80	3.52	11 (5 - 16)	8.50	4.93	11 (3-14)	8.52	4.23	16 (2-18)	6.07	4.23	16 (0 – 16)
Sexual Offences	6.00	3.50	9 (3 - 12)	3.50	.58	1 (3-4)	4.66	2.87	12 (0-12)	3.55	2.71	11 (0-11)
Future Planning	2.80	1.32	3 (1-4)	2.25	1.71	4 (0-4)	2.59	1.51	6 (0-6)	2.27	1.56	24 (1-25)
SVR-20 Overall	18.70	6.50	20 (12-32)	14.25	6.60	15 (7-22)	15.84	6.56	29 (3-32)	11.93	6.24	4 (0-4)
ARMIDILO-Stable	8.50	5.06	19 (-3 -16)	7.50	6.19	14 (-1 -13)	7.22	5.83	27 (-11-16)	3.48	5.43	21 (-7 -14)
ARMIDILO-Acute	.80	2.44	8 (-2-6)	3.00	3.40	8 (-2 – 6)	1.07	2.78	14 (-7-7)	.36	1.64	8 (-4 -4)

Note. Special needs group combined both ID offenders, borderline functioning offenders and those offenders without formal IQ assessments.

Table 2.3: Predictive Validity of the Risk Instruments used in the Study

Recidivism		Special Needs Group (n = 44)				Non-ID Group (n = 44)				Total Sample (n = 88)			
		Unofficial[a]	Sexual	Violent	General	Unofficial[a]	Sexual	Violent	General	Unofficial[a]	Sexual	Violent	General
RRASOR		.52	.54	.70	.65	.63	.60	.55	.50	.57	.57	.65	.58
RM2000-V		.48	.50	.57	.82	.45	.56	.71	.83	.49	.53	.63	.82
SVR-20	Psychosocial Adjustment	.42	.50	.81	.53	.70	.72	.48	.65	.57	.60	.73	.59
	Sexual Offences	.42	.37	.55	.46	.62	.66	.63	.45	.55	.54	.62	.47
	Future plans	.55	.50	.77	.59	.44	.47	.63	.56	.51	.48	.73	.59
	Overall score	.40	.45	.80	.50	.70	.73	.55	.60	.58	.59	.74	.56
ARMIDILO- Stable		.56	.61	.83	.62	.64	.63	.62	.57	.62	.61	.79	.61
ARMIDILO- Acute		.76	.73	.76	.59	.48	.40	.52	.50	.66	.57	.72	.56

Note. [a] Includes unofficial sexual recidivism data and reconviction data.
AUC = Area under the curve. All AUC of the ROC calculations were performed using an open-ended, total follow-up time frame. A 5-year follow-up period was used. Special needs group consisted of offenders assessed as borderline functioning, those with intellectual disabilities and offenders without formal IQ assessments.

Table 2.4: Predictive Validity of the Risk Instruments for Special Needs Subgroups

Scales		ID (IQ< 75) Group (n=10)			Borderline (IQ>75) Group (n=4)		
		AUC	*SE*	95% CI	AUC	*SE*	95% CI
RRASOR		.47	.24	-.00 to .94	.50	.35	-.19 to 1.19
RM2000-V		.63	.21	.21 to 1.04	1.0	.00	1.00 to 1.00
SVR-20	Psychosocial Adjustment	.88	.12	.65 to 1.1	.50	.35	-1.19 to 1.19
	Sexual Offences	.41	.24	-.06 to .87	.00	.00	.00 to .00
	Future plans	.66	.19	.28 to 1.03	.50	.35	1.19 to 1.19
	Overall score	.75	.15	.45 to .1.05	.50	.35	-1.19 to 1.19
ARMIDILO- Stable		.86	.13	.62 to .1.13	.75	.28	.20 to 1.30
ARMIDILO- Acute		.75	.15	.45 to .1.05	.50	.35	-.19 to 1.19

Note. Based on sexual reconviction data. Only 14 offenders within the special needs group had received formal IQ assessments .
AUC = Area under the curve. All AUC of the ROC calculations were performed using an open-ended, total follow-up time frame. A 5-year follow-up period was used

It can be seen from Table 2.3 that for the special needs group the ARMIDILO-Acute instrument was the best predictor for sexual reconviction. For predicting unofficial recidivism and sexual reconviction after a five year follow-up period, the ARMIDILO-Acute produced AUCs of .76 (95% confidence interval [CI] .61 to .91) and .73 (95% confidence interval [CI] .57 to .89) respectively, which were significantly above chance level ($p < .001$). While, RRASOR performed little better than chance level for predicting sexual recidivism in the special needs group.

In terms of violent recidivism the SVR-20 Psychosocial, Future Planning and Overall scales were good at predicting violent reconviction, but only for the special needs group. However, the RM2000-V was better able to predict general recidivism than violent recidivism despite being designed to predict violence, as it significantly predicted general recidivism for both the special needs and the non-ID groups, with AUCs of .82 and .83 respectively.

Within the special needs group, offenders with an intellectual disability were analysed separately (see Table 2.4), based on a small sample of 10 participants. The ARMIDILO-Stable scale produced a significant predictive effect, with AUCs of .86 (95% CI .62 to 1.13). The ARMIDILO-Acute, SVR-20 Psychosocial Affect and Overall scales were also better predictors of sexual recidivism for the ID group, achieving AUCs of .75 (95% CI .45 to 1.05), .88 (95% CI .65 to 1.10) and .75 (95% CI .45 to 1.05) respectively. It can be seen from Table 2.4 that the RRASOR performed little better than chance for predicting sexual recidivism within the special needs group, although almost produced a moderate effect size for the non-ID sample. There was however no statistically significant difference between both the ID and the combined special needs and mainstream groups on this scale.

In a secondary analysis to ascertain whether any of the risk assessment instruments were significantly better at predicting sexual reconviction within the special needs group, a series of paired z-score were performed. Comparisons indicated no significant differences in the predictive accuracy of the RRASOR, RM2000-V, SVR-20 or the ARMIDILO scales. Although the ARMIDILO-Acute scale achieved the highest AUC for unofficial and official sexual recidivism, it was not however significantly better than the other risk instruments.

For comparative purposes, the reliability estimates of the risk assessment measures were calculated using Pearson *r* coefficients between risk scores and the reconviction outcome, to determine whether a relationship existed. Reoffence type

was separated between unofficial and official sexual, violent and general recidivism. Discrepancies between the tests were noted. For the non-ID sample with regards to sexual recidivism the SVR-20 Overall scale established a significant correlation (r = .33, p < .05) as well as the SVR-20 Psychosocial scale (r = .36, p < .05). All the measures failed to show significant correlations between sexual reconviction for the special needs group. For the special needs group, only the ARMIDILO-Acute significantly correlated with unofficial sexual recidivism (r = .46, p < .01). Three scales significantly correlated with violent recidivism for the special needs sample (SRV-20 Psychosocial Adjustment, r = .35, p <.05; SVR-20 Overall, r = .33, p <.05 and ARMIDILO-Stable, r = .32, p <.05). The RM2000-V did not correlate with violent recidivism for either the special needs or non-ID groups. For general recidivism the RM2000-V significantly correlated with for the special needs group (r = .52, p <.01) and the non-ID group (r = .55, p <.01).

Discussion

The primary purpose of this study was to access and compare the predictive accuracy of risk assessment instruments in predicting recidivism with special need sexual offenders. Only the ARMIDILO instrument was the best predictor for sexual reconviction for the special needs group. It was found that the risk assessment tools used within this study were more accurate predictors of sexual recidivism for the non-ID group than for the special needs group. With the exception of the ARMIDILO tool, risk assessment instruments in this study failed to assist in estimating risk levels in the special needs sample. Unfortunately, the RRASOR and the SVR-20 scales obtained lower AUC indices and performed little better than chance level at distinguishing sexual recidivists from non-recidivists in the special needs group. Unexpectedly, the SVR-20 actually performed worse than chance for the special needs sample, although did produce moderate predictive validity with the non-ID sample. It had been predicted that the SVR-20 may have performed better than the RRASOR given that the SRV-20 contained a number of relevant variables. The special needs sample yielded significantly higher risk ratings on the SVR-20 Psychosocial Adjustment and Overall scales, and generally obtained higher risk scores on all of the instruments used within this study, indicating that risk scales over-predicted the level of risk for this population of sexual offenders. The results also indicated that the ARMIDILO-Stable

scale produced a similar AUC for predicting sexual recidivism for both the special needs and non-ID groups, suggesting that the factors within this scale did not contribute uniquely to the prediction of sexual recidivism with offenders with special needs, these risk variables appear to be equally as relevant to both samples. The ARMIDILO-Acute scale showed higher predictive validity for the special needs offenders. A possible explanation for this could be that these offenders may not have mastered the skilled in the art of deception, being perhaps less able to mask changes in their emotional states, coping abilities and sexual preoccupations compared to those without intellectual impairments and so obtained higher scores on items within this scale.

In terms of violent recidivism, for the special needs group the SVR-20 Psychosocial Adjustment, Future Planning and Overall scores obtained a moderate predictive accuracy for violent reconviction. Interestingly, the RM2000-V scale did not predict violent recidivism as well as the SVR-20 scales for this group. However, for the non-ID group the RM2000-V was able to better predict violent recidivism compared to other instruments. Within the special needs group, for the intellectually disabled subgroup the ARMIDILO-Acute, SVR-20 Psychosocial Affect and Overall scales were good predictors of sexual recidivism.

In accordance with previous research by Wilcox et al. (2009), these findings are consistent with the position that actuarial risk tools are not as effective for intellectually disabled populations. Although the RRASOR yielded a higher AUC in this study than in the Wilcox study, such a difference may be attributed to the larger sample size incorporated in this study. Wilcox et al. (2009) recognised that some of the actuarial risk instruments contained biased items that predisposed intellectually disabled offenders to have a heightened level of risk category (such as long-term relationship status). Wilcox argued that this would increase the level of risk as intellectually disabled individuals are less able to maintain lasting relationships. The SVR-20 and the ARMIDILO tool both do not contain a time restraint (i.e. having lived with an adult lover for at least two years, as referenced in the RM2000-S instrument), but rather include a broader definition of ‘relationship problems’ or ‘relationship skills’ to increase the level of risk.

This study showed that risk assessment instruments incorporating both acute and stable dynamic risk domains appear to be a more accurate assessment method for a population with intellectual impairments. This research also provided support for

Boer, Tough and Haaven's (2004) assertion that additional dynamic assessment would heighten the accuracy of the risk assessment for intellectually disabled sexual offenders.

Limitations of the study

The primary limitation of this study was the retrospective follow-up design, as this increased the likelihood of predictor-criterion contamination. The rater was not masked to the recidivism status, given that sexual recidivism information was contained within the file notes. The rater was, however, blind to the reconviction status as official reconviction data was not collected until the risk assessments had been completed. File-coding also increased the likelihood of an underestimation of scores for the measures, as it was subject to retrospective recall biases, and bias in case note information recorded within the files. Furthermore, inter-rater reliability could not be reported as this study relied on only one rater.

In order to counteract the problems with official reconviction sources, this study utilised two forms of outcome sources to detect sexual recidivism (Home Office and unofficial data) in order to avoid an underestimation of risk. Unfortunately, information regarding sexual reoffending was not available or recorded in the Probation files, and sexual recidivism data was limited to only the time in which the offenders were on Probation Orders and receiving treatment. When applying dichotomous binary distinctions (i.e. sexual or non-sexual reconviction), valuable information was also overlooked, as the ROC curve analysis does not reflect the frequency or the severity of reoffending, thus the Harris and Rice (2003) argue that the predictive accuracy of actuarial scales are limited by a methodological ceiling effect.

It should also be noted that the method of assessing intellectual functioning in this study was not consistent, as some participants were given the WAIS and others the SILS. Conclusions regarding the intellectually disabled offender group were based on a small number of participants. Only 14 participants actually received an IQ assessment, four of whom were within the borderline level of functioning range, and so interpretations regarding the analysis of the borderline subgroup were not possible within this study given such a small number of participants.

The SILS is a brief screen measure and is not a comprehensive multidimensional test of intelligence. It is heavily reliant upon verbal skills that tend

to be acquired through formal education. The accuracy of the SILS in estimating WAIS-R Full Scale IQ scores has been questioned. Frisch and Jessop (1989) argued that failing to control for reading level may lead to an underestimation of IQ scores. Reading level was not controlled for within this study and higher IQ scores may be underestimated, and so the treatment group may have comprised men with greatly varying intellectual levels, thus skewing the data. Future studies are recommended to assess IQ in a standardised and uniform fashion as this would allow for comparisons between borderline functioning and intellectually disabled offenders.

Suggestions for future research

Replication of this study using a larger sample size, based on a prospective design to confirm the results found within this study would be of value. The inclusion of interview-based sources of information, and the use of psychometric testing would also increase the reliability of the data. While the preliminary findings of the ARMIDILO are promising, they are tentative and require further validation in order to support its use in clinical and forensic practice.

In light of the findings of previous research, it would appear that the use of structured risk assessment tools are not as accurate when applied to offenders with intellectual impairments as to those without. This raises ethical questions as to whether such instruments should continue to be used with this population. The use of ineffective instruments could have negative implications on decision-making and the risk management of such offenders. Further research is warranted to confirm the findings of this research.

Conclusions

This research is hopefully of practical value, as it adds to the empirical knowledge regarding the predictive utility of risk assessment instruments and helps inform psychologists when assessing the level of risk in those sex offenders who have intellectual impairments. This study questioned the applicability of a wide range of risk assessment tools to this group of sexual offenders.

This research also highlights the importance of including dynamic risk factors in the assessment of risk. The ARMIDILO, although in its very early stages of development, was designed to address risk factors specific to intellectually disabled

sexual offenders. However, further empirical validation of this tool is required. Given that dynamic factors were a better indicator of risk levels with this population, it may be more fruitful for practitioners to target treatment to the idiosyncratic dynamic risk factors for this specific population in order to reduce future recidivism.

Chapter Four

A Critical Appraisal of a Risk Assessment Instrument: the RRASOR

It is paramount that accurate risk assessments be employed. The RRASOR (Hanson, 1997) a purely static risk assessment tool widely used by clinicians, has been subject to much empirical validation, as identified in the first chapter, and this was used in the research outlined in the second chapter. This chapter offers an overview of the tool and presents an in depth critical appraisal of this static risk instrument in terms of its consistency over time and the strengths of its inferences. The concurrent, construct, content, discriminant validity, reliability and standardization of the RRASOR will be examined. The chapter then explores some of the difficulties associated with using the RRASOR.

Introduction

The assessment of risk is a complex task which is ethically challenging and requires optimal precision (Craig, Browne, Stringer & Beech, 2005). The qualities of a good risk assessment instrument incorporate appropriate standardisation of the test, along with high levels of predictive validity and rater reliability, and the inclusion of important risk factors (Hanson, 1998; Campbell, 2000), and manual availability. Sufficient peer review is also necessary for evaluation of risk assessments (Campbell, 2000). This review examines the Rapid Risk Assessment for Sexual Offence Recidivism (RRASOR; Hanson, 1997). An overview and critique of the instrument will be provided and assessed in terms of its validity, reliability and clinical utility.

The RRASOR is an actuarial tool containing four static historical variables. These items are *past sexual offences, age at commencement of risk, extra familial victims,* and *male victims.* The RRASOR provides a probabilistic estimate of risk, using percentages of people with the same score who would be expected to reoffend within a specified period of time. Offenders are allocated points according to the presence of these items and a given risk categorisation is provided on the basis of a score ranging from 0-6, with the higher scores incrementally associated with an increased risk for recidivism. Although there are only four items, for the first item 'prior sex offences' offenders can be given a score of 2 for having 2-3 convictions and a score of 3 when they have 4 or more convictions, they then get an additional point for being younger than 25 years old, having any male victims or any unrelated victims. The corresponding recidivism rates over a 10 year period of opportunity range from 6.5 to 73.1%.

The RRASOR was developed in Canada from a predominantly North American male sample (mean age varied between the samples from 30.4 to 37.6 years). It was derived initially from a development sample which consisted of seven different follow-up studies, which were then cross-validated using a separate independent sample. The developmental samples represented a variety of settings in which risk assessments are usually conducted, these included prisons, specialised treatment programmes, and secure mental health facilities. The validation sample was based on a British prison setting, and was selected on the basis of a long follow-up period of 16 years, and a sufficient sample size (n = 303). The overall sample size was very large (n = 2592) when compared to other scales, and is above Hanson (1998)

recommended sample size of 1000 offenders for a study to adequately establish the validity of an assessment procedure.

Items for inclusion in the actuarial instrument were selected from an initial pool of seven predictor variables from Hanson and Bussière's (1996) meta-analysis. The scale was then constructed based on the results of a stepwise regression analysis, which determined the best four risk predictor variables.

1. Validity

Predictive Validity

The predictive accuracy of a risk assessment tool is best indexed through the Area Under the Curve statistic (AUC) of the Relative Operating Characteristics (ROC) analysis (Rice & Harris, 1995). The AUC result can be interpreted as the probability that a randomly selected recidivist would have a higher score on a risk instrument than a randomly selected non-recidivist. An AUC value of .50 represents a chance prediction, whereas an AUC of 1.0 represents a perfect positive prediction. In general an AUC value of .56 corresponds to Cohen's *d* value of .2, representing a small effect. An AUC of .65 is equivalent to a Cohen's *d* value of .5, reflecting a moderate effect, whereas an AUC value equal to and above .70 is comparable to Cohen's *d* value of .8, which can be considered a large effect (Rice & Harris, 2005).

Hanson (1997) reported moderate levels of predictive accuracy of the RRASOR (AUC = .71, *r* =.27) based on all samples used in the test construction. The AUC values ranged from .62 to .77. There was however, no significant variability in the AUC values reported across the seven developmental samples and the validation samples. Since the development of the RRASOR a number of studies based on large sample sizes have gone on to test the predictive accuracy of this actuarial instrument. It can be seen from Table 3.1 that the AUC values for sexual recidivism ranged from .59 to .77, with only two studies (Harris, Rice, Quinsey, Lalumière, Boer & Lang, 2003; Bartosh, Garby, Lewis & Gray, 2003) yielding a RRASOR with a small effect. All other studies produced moderate to large effects in the predictive accuracy of the instrument, suggesting that the predictive validity is relatively consistent and is supported by a number of empirical studies. Furthermore, research also indicated that the RRASOR has greater accuracy in the prediction of sexual recidivism in comparison to the prediction of violent and general recidivism.

Table 3.1. Predictive validity of the RRASOR

Author(s)	Sample Size	Sample Origin	Recidivism		
			Sexual AUC	Violent AUC	General AUC
Hanson and Thornton (2000)	1225	Canada & UK	0.68 *(r*=.28)		
Barbaree et al. (2001)	215	Canada	0.77 *(r*=.26)	0.65* *(r*=.20)	0.60 *(r*=.14)
Sjöstedt and Långström (2001)	1384	Sweden	0.72 *(r*=.22)	0.63 (*r*=.17)	
Harris et al., (2003)	396	Canada	0.59	0.56	
Bartosh et al., (2003)	186	USA	0.632 *(r*=.15)	0.57 (*r*=.094)	0.597 (*r*= .12)
Långström (2004)	1303	Sweden	0.73	0.59	
Seto et al., (2004)	113	Canada	.83	.68	
Langton et al., (2007)	468	Canada	.68		*.57*

NOTE: Recidivism was recorded as a dichotomous outcome, i.e. 'yes or no'. Samples are based on mainstream sexual offenders. * Barbaree et al. (2001) referred to serious recidivism rather than violent recidivism.

The studies in Table 3.2 show that the predictive validity does not generalise across offender ethnicity or migrant status (Långström, 2004) and should not be used with Indigenous offenders in Australia without further research (Allan, Dawson & Allan, 2006). The RRASOR is also not consistent with regards to the levels of risk in offender types and is of no value at discriminating non-contact recidivists from non-recidivists (Bartosh et al., 2003).

Similarly, Craig, Browne & Stringer (2004) found that those with child victims obtained significantly higher scores on the RRASOR. The predictive validity was also found to vary when the criteria variable was elaborated, for instance, Sjöstedt and Grann (2002) found that the RRASOR was found to better predict imminent and less severe reoffending than repeated and injurious sexual re-offences.

Despite limitations in application to certain types of offenders, it has been suggested that the RRASOR may provide a useful metric of risk for the intellectually disabled population (Harris & Tough, 2004; Tough, 2001). However, recent data

based on a sample of intellectually disabled sexual offenders yielded a very low AUC of 0.42, indicating that the RRASOR performed little better than chance at predicting sexual recidivism (Wilcox, Beech, Markall & Blacker, 2009). The exclusion of self-reported sexual offences in the scoring of the instruments may hinder the reliability of the tool, given that sexual offenders with intellectual disabilities are less likely to be prosecuted for sexual offences (Swanson & Garwick, 1990; Thompson & Brown, 1997). The predictive accuracy of the RRASOR may increase for this population if sexually abusive behaviours, not simply charges and convictions are taken into account.

Table 3.2. Predictive validity of offender characteristics using RRASOR.

Author(s)	**Sample Size**	**Offender/Offence type**	**Sexual Recidivism AUC**
Wilcox et al. (2009)	27	Intellectual Disabilities	.42
Långström (2004)	128	African Asian	.48
	49	European	.77
	1085	Nordic	.76
Bartosh et al., (2003)	17	Non-contact	.49
	73	Rape	.53
	37	Incest	.73
	59	Extra-familial	.58
Sjöstedt & Grann (2002)	1288	Extra-familial	.81
		Intra-familial	.44
Sjöstedt & Långström (2002)	51	Rapists	.73

In a meta-analysis of studies the Q statistic is a measure of consistency, testing the homogeneity of effects between studies and used to compare the variability of predictive accuracy across different samples. A significant Q statistic indicates that there is more variability across studies than would be expected by chance, which could not be attributed to either variability in scoring procedures or differential validity across samples, rather than a single outlier study. The most reliable findings are those with low variability across studies. Hanson and Thornton (2000) found the RRASOR to be consistent across four diverse prison and secure psychiatric samples with a Q of 3.56, which is no greater than would be expected by chance. By contrast

in a recent meta-analysis Hanson and Morton-Bourgon (2004) identified significant variability in the prediction of sexual recidivism across 18 studies, with a Q statistic of 55.84, suggesting that the scale does not appear to be very reliable, although no explanation has been provided to explain such variability in the findings.

In accordance with the "Ethical Standards" and "Testing Standards" set by the American Psychological Association (1992), psychologists engaged in risk assessment and professional test developers are required to address or determine four prediction outcomes associated with the risk assessment instrument; 1) What is the level of *sensitivity* or *hit rate* of the assessment procedure used for the risk assessment (percentage of previously convicted offenders who will recidivate are correctly predicted as high-risk by the assessment procedure). 2) What is the level of *specificity* of the assessment procedure used for risk assessment (percentage of previously convicted offenders who will not recidivate are correctly identified as low-risk by this assessment procedure), and two sets of errors; 3) What is the frequency of the *false positive* or *false alarms* (nonrecidivists incorrectly predicted to recidivate) classifications associated with this assessment procedure?, and 4) What is the frequency of *false negative* or *misses* (recidivists incorrectly deemed to not recidivate) classifications associated with this procedure?

When defining the classification accuracy and evaluating of the quality of the actuarial risk assessments it is therefore necessary to report the frequencies of "hits" and "misses" associated with given cut-off scores, so that conclusions can be made regarding whether an offender warrants commitment under the relevant statue (Campbell, 2000). The miss rate (false negative) is reciprocal to specificity (true negatives), whereas the false alarm rate (false positive) is reciprocal to sensitivity (true positive).

With regards to the RRASOR Hanson (1997) omitted to report the greatest level of overall classification accuracy. The levels of sensitivity and specificity were not reported, or the frequencies for the false alarm and the miss rates provided. However, Janus and Meehl (1997) proposed that if a maximum cut-off point of 5 was implemented to predict recidivism for legal commitments, the RRASOR would yield very unequal levels of sensitivity and specificity, with a specificity score of .97 and a low sensitivity score of .15. This implies that 85% of offenders would be mistakenly classified as non-recidivists and would go free into the community, and only a mere 15% would be correctly identified as recidivists, and so therefore should not be used

in legal proceedings. Similarly, in a study by Sjöstedt and Långström (2001), they reported an optimum cut-off score of >2 for the RRASOR, with a sensitivity of 0.37 and a specificity score of 0.96. Campbell (2003) pointed out that this would mean that the cut-off misses an estimated 63% of the previously convicted offenders assumed to recidivate if released from prison and this would thus hinder public protection. By adjusting the cut-off score to maximise the sensitivity levels of the RRASOR this involves increasing the false positive classification and significantly reduces the specificity. With such disparity between the levels of sensitivity and specificity, the validity is low and it is argued that this procedure should not be used in isolation. With the lack of data regarding the levels of sensitivity, specificity and the frequencies of false positives and false negatives associated with the RRASOR, the RRASOR has not adhered to the ethical and testing standards placed by the American Psychological Association (1992), which compromises its use or in legal practice (Campbell, 2000).

Concurrent Validity

Concurrent validity studies using the RRASOR have been uniformly supportive, indicating that the RRASOR correlates well with other risk assessments of sexual recidivism (Barbaree, Seto, Langton & Peacock, 2001; Craig, Browne & Stringer, 2004; Roberts, Doren & Thornton, 2002). Barbaree et al. (2001) found the RRASOR was highly correlated with the Static-99 (r =.75), compared with other risk instruments such as the MnSOST-R (r =.32), the SORAG (r =.38), and the VRAG (r =.14), however, the RRASOR was not significantly correlated with the Psychopathy Checklist-Revised (r =.13). The RRASOR was found to be the most accurate (AUC=.77), whereas the AUCs for the Static-99 and the SORAG were both .70, with the VRAG and the PCL-R the least accurate (AUC=.61). Similarly, Langton et al., (2007) also reported that the RRASOR was highly correlated with the Static-99 (r =.77) and the Static-2002 (r =.69), although the Static-2002 was significantly more accurate in predicting sexual recidivism than the RRASOR.

More specifically in comparison to other risk assessment tools, Hanson (1997) noted that the level of predictive accuracy is higher than that found using more elaborate scales such as the VRAG as Rice and Harris (1997) found that the VRAG has an AUC of .62 for predicting sexual recidivism using a sample of child molesters and rapists. Similarly, Barbaree et al. (2001) noted that the RRASOR outperformed

the VRAG, Static-99, SORAG and the MnSOST-R (see Table 3.3). However, Harris, Rice, Quinsey, Lalumière, Boer and Lang (2003) compared the RRASOR with the VRAG, SORAG and the Static-99, and found that the RRASOR had lower ROC area than the other risk assessments. Similarly, from the AUC comparisons across different risk instruments in Table 3.3, it can be seen that the predictive accuracy was not as strong as the Static-99 and SORAG when predicting sexual recidivism for rapists (Bartosh et al., 2003). However, in the Hanson and Morton-Bourgon (2004) meta-analysis, no significant differences were found among specific risk assessment tools in predicting sexual recidivism.

Table 3.3. ROC score comparison for different actuarial risk instruments

Author		**RRASOR**	**Static-99**	**MnSOST-R**	**SORAG**	**VRAG**
Barbaree et al., (2001)		.77	.70	.65	.70	.61
Bartosh et al., (2003)	General	.63	.64	.59	.58	-
	Rapist	.53	.71	.54	.71	-
	Intra-familial	.73	.74	.63	.72	-
	Extra-familial	.58	.65	.59	.70	-
Harris et al., (2003)	Child molester	.61	.65	-	.70	.70
	Rapist	.56	.59	-	.62	.64

Note: AUC's for sexual recidivism only. Static-99(Hanson & Thornton, 2000); MnSOST-R (Minnesota Sex Offender Screening Tool–Revised; Epperson, Kaul, Huot, Hesselton, Alexander & Goldman, 1998); SORAG (Sex Offender Risk Appraisal Guide; Quinsey, Harris, Rice & Cormier, 1998); and VRAG (Violence Risk Appraisal Guide; Webster et al., 1994).

Construct Validity

The construct validity of the RRASOR was touched upon by Barbaree, Langton and Peacock (2006). They found that the RRASOR loaded almost exclusively with sexual deviance factors (persistent, child sexual abuse, male victims, detached predatory behaviour) and concluded that it was therefore primarily an assessment of sexual deviance rather than anti-social behaviour. Persistence and child sexual abuse were significant predictors of sexual recidivism. Previously sexually deviant interest and anti-social orientation were found to be the two most important groups of predictor variables for sexual reoffending (Hanson & Morton-Bourgon, 2004).

Content Validity

During the construction of the RRASOR an initial pool of predictor variables was selected from Hanson and Bussière's (1996) meta-analysis. Using a step-wise regression, variables selected for inclusion were those that had an average correlation of at least .10 with sexual offence recidivism, and those that could be scored using commonly available information. A weakness of the RRASOR identified by Hanson (1997) was that deviant sexual preference was found to be the strongest recidivism predictor in the meta-analysis, yet the RRASOR fails to directly consider sexual preference, thus diminishing the RRASOR's content validity.

Discriminant Validity

The RRASOR has consistently been shown to be better at predicting sexual recidivism over both violent and general recidivism and distinguishing sexually violent from general violent recidivist participant. Both Sjöstedt and Långström (2001) and Långström (2004) found AUCs of .72 and .73 respectively for sexual recidivism, but much lower AUCs for violent recidivism, with AUC of .63 and .59 (see Table 3.1). Similarly, Bartosh et al. (2003) found that the RRASOR was significantly predictive of sexual re-offenders, but not of offenders who only recidivated violently. This suggests that the RRASOR does have good discriminant validity

2. Reliability

There are two different types of reliability that pertain to risk assessment tools: *inter-rater* and *intra-rater* reliability. Inter-rater reliability refers to two or more people scoring the same rating for an individual, whereas intra-rater reliability involves a single rater completing the assessment for the same offender on repeated applications, with the differences in rating then compared. With intra-rater reliability the assessors subsequent ratings are contaminated by knowledge of earlier ratings, and so inter-rater reliability tends to be used more. Sufficient levels of inter-rater reliability are necessary for a risk assessment procedure to be standardised, poor reliability ultimately compromises the validity of the assessment (Campbell, 2000).

Reliability interpretations were not included in the original manual by Hanson (1997). However, studies have shown that the RRASOR generally holds good inter-rater reliability indexes (Austin, Peyton & Johnson, 2003; Barbaree et al., 2001;

Harris et al., 2003; Sjöstedt & Långström, 2001, 2002; Långström, 2004), indicating that the RRASOR's reliability does not compromise the validity of the tool, as it can be applied consistently. Furthermore, Doren (2004) found that the five year interpretative risk percentages for the RRASOR were replicated and were remarkably stable despite changes in the sample's underlying recidivism base rates. However, Austin, Peyton and Johnson (2003) argued that significant staff training and item refinement was needed to ensure consistent scoring across raters. To date no studies have examined intra-rater reliability.

3. Clinical utility

The RRASOR is quick and easy to complete using file information, simplicity therefore is one of its greatest virtues. It may also be useful in assessing violent recidivism (Bartosh et al., 2003). However, there is no generally available manual for the RRASOR detailing its proper use. Also, the RRASOR is not appropriate to be used in isolation based on the limited information contained within the scale. It was intended as a screening measure of the relative risk levels, and not as a comprehensive assessment of all the factors relevant to the prediction of sexual offender recidivism. Consideration of dynamic risk factors, treatment cooperation and other relevant information, such as the level of community supervision can thus be included and the risk level adjusted accordingly (Hanson, 1997).

It terms of its application Sjöstedt and Långström (2001) argue that the RRASOR needs to be supplemented with other non-actuarial approaches and advise that extreme caution should be exercised when applying the RRASOR outside of a research framework. It has further been urged against using the RRASOR within the legal context, for instance for expert testimony in a legal proceeding (Sjöstedt & Långström, 2001; Campbell, 2000). Janus and Meehl (1997) argued that the actuarial assessment should demonstrate accuracy rates of .70 or .75 for the judicial system to make valid and reliable judgements on civil commitment laws, unfortunately, AUC rates of .70 or greater have not always been consistently found across different samples.

Austin, Peyton and Johnson (2003) examined recidivism rates of 550 sexual offenders and found that recidivism dropped dramatically in offenders aged 45 and older, whereas those in all other age groups the recidivism rates remained fairly consistent. Given the age variable on the RRASOR is based on a simple scoring

dichotomy (age 25 years or older versus under 25 years of age) yet very few offenders fall under the age 25 designation when they are eligible for release from prison, this would have little impact on the assessment process. Austin, Peyton and Johnson (2003) therefore recommend that the level of predictions associated with the RRASOR could be improved by adjusting the age cut-off threshold level from age 25 to 45 years at the time of release.

In terms of risk classification and labelling, the RRASOR has provided clinicians and researchers with a useful metric assessment of risk of sexually reoffending and has contributed to the area of sexual risk assessment. However, the RRASOR is subject to the limitations described in the actuarial risk assessments. For instance, Beech, Fisher and Thornton (2003) criticised the actuarial approach for excluding dynamic items, focusing exclusively on static risk factors that are unchangeable and cannot be addressed in treatment. Furthermore, Beech, Fisher and Thornton (2003) argue that such risk assessments are based on underestimations of true offence rates, thus increasing the potential for false negative error rates. Moreover the studies described in this critical appraisal have used reconviction over a fixed time period as an outcome measure which does not take account of the severity of reoffending.

4. Conclusions

The RRASOR has been used in various forensic settings, and has good empirical foundation, with a number of merits, including its large test construction sample, and its time efficiency, as it can easily be scored from administrative records. It has also achieved considerable research interest over the past decade, with numerous cross-validation studies, reflecting its general acceptance via peer review. A review of the literature suggests that the RRASOR holds moderate and consistent levels of predictive validity for the general sexual offender population and has yielded good inter-rater reliability across a number of studies.

Whilst the RRASOR was standardised on a large sample thus demonstrates strong generalisability of the findings across different forensic settings and countries. The RRASOR does however exclude the assessment of juvenile, female sexual offenders and others from other non-western cultures. Caution is also advised when applying the RRASOR to offenders with different ethnicities (Långström, 2004;

Allan, Dawson & Allan, 2006) and to offenders with intellectual disabilities (Wilcox et al., 2009).

The RRASOR further contains a number of inherent weaknesses. The RRASOR does appear to exclude many risk factors considered to be important to sexual recidivism, such as deviant sexual preference. It is limited by focusing exclusively on static risk factors in the assessment of risk, thus it does not allow change to be measured, or reflect fluctuations in risk. Furthermore, the specificity and sensitivity, the false positive and false negative rates were not commented on by the author. Finally, the RRASOR cannot support expert testimony in legal proceedings, although Hanson (1997) does acknowledge that the RRASOR should not be used in isolation and did recommend that it be used only to screen offenders into the relative risk levels, rather than providing a comprehensive assessment of risk factors relevant to the prediction of sexual recidivism. It is therefore necessary that both researchers and clinicians continue to scrutinize the application of this tool given the limitations discussed.

Chapter Five

Risk Assessment in Clinical Practice: Assessment and Intervention of a Special Needs Sexual Offender

Identifying the level of risk and the factors linked with sexual recidivism is imperative to the understanding of appropriate interventions for the prevention of further sexual abuse (Craig, Beech & Browne, 2006b; Craig, Browne, Stringer & Hogue, 2007). The previous chapter examined the application of static and dynamic risk assessment tools to sexual offenders with intellectual disabilities. This research found that risk assessment tools incorporating dynamic risk variables were more effective for this population in comparison to mainstream offenders and that the predictive validity was higher for tools incorporating dynamic risk factors. With static risk instruments any positive gains in treatment will not be recognised as treatment cannot change static risk variables. However, identifying dynamic risk factors associated with offending behaviour can direct clinicians as to the treatment needs and goals.

Recently Craig, Browne, Hogue and Stringer (2004) provided a Multiaxial Risk Appraisal (MARA) Model that encourages a global procedure to the assessment of risk by considering the "nomothetic" approach. This approach involves the use of actuarial scales and psychometric assessment of psychopathology and psychosexual characteristics, and the 'idiographic' domains, which rely on dynamic changes in risk and empirically guided clinical judgement. The idiographic approach is based upon person centred factors of risk which are unique to the individual, whereas the nomothetic approach compares risk to large group norms.

The objective of this case study is to show the risk assessment process based on the Multiaxial Risk Appraisal Model to assess risk of future sexual violence, and to utilise structured professional checklists, i.e. the SVR-20 (Boer et al., 1997) and the ARMIDILO (Boer, Tough & Haaven, 2004) to help formulate the treatment goals in order to reduce the level of risk in a sexual offender with an intellectual impairment.

An assessment and functional analysis indicated that social and interpersonal skills deficit was evident, and that a social skills intervention was required to address this. Relationship skills problems are a stable dynamic risk variable incorporated into both the SVR-20 and the ARMIDILO risk instrument. It has been recognised that sexual offenders with intellectual disabilities have poor peer relations (Hayes, 1991)

an inability to understand normal sexual relationships, and lack relationship skills (Day, 1994). It has been argued that interpersonal skills are a necessary component that needs to be integrated into treatment program for intellectually disabled sexual offenders (Coleman & Haaven, 2001).

This chapter additionally aimed to evaluate a social skills training intervention and highlight some of the difficulties associated with working therapeutically with intellectually disabled individuals.

Abstract

Social skills deficits have been associated with sexual offending behaviour (Marshall, 2001). Poor social skills are particularly prevalent in the intellectually disabled sex offender population, who are characterised by limited relationship skills, high levels of loneliness, low self-esteem (Boer et al., 1995) and poor social functioning (Griffiths, Hindsburger & Christian, 1985).

This case study is based on a single case (A:B:A) design assessment and intervention for an intellectually disabled sexual offender, who had been given a two year Probation Order for indecent assault against a fifteen year old girl. He had completed a one year Adapted Sex Offender Treatment Programme, however facilitators expressed concerns that he had not benefited from treatment and that his level of risk remained the same. A referral was subsequently made for individual treatment.

Assessment and functional analysis resulted in the formulation that the client would benefit from a social skills intervention. A social skills programme was designed specifically for the client, and was administered in an individual rather than group format. The treatment consisted of ten, two-hour sessions, aimed at increasing the client's social skills with particular emphasis on enhancing eye contact and assertiveness. The intervention incorporated a combination of techniques, such as role-play and the introduction of concepts of 'old-me, new-me'.

Pre-post treatment outcome measures showed no significant treatment gains. Self-disclosures also indicated that the client had difficulties in retaining the information during the sessions. However, the client had partially benefited from the intervention in terms of improvements with regards to the level of eye contact. The outcome of this case study is discussed in relation to the intervention setting, client characteristics and assessment issues. Therapist dependence issues were also a concern throughout the intervention.

Introduction

Sexual offending by men with intellectual disabilities has only been recognised in the past decade as an important area for research (Rose, Jenkins, O'Connor, Jones & Felce, 2002; Wilcox, 2004; Thompson & Brown, 1997). Treatment programmes designed especially for intellectually disabled sex offenders have just been established, in order to make treatment available for those offenders who are not suitable for mainstream sex offender programmes (Keeling, Rose & Beech, 2006). Intellectual disability was often an exclusion criterion used in many treatment programmes (O'Connor, 1997), and therefore, many have received inadequate services for their sexual offending behaviour (Lambrick & Glaser, 2004). Treatment resources for this population have also generally been inadequate (Rose et al, 2002). Research on the specialised treatment of intellectually disabled sex offenders has therefore been slow in comparison to 'mainstream' sex offender treatment (Wilcox, 2004). Thompson and Brown (1997) argued that clinical interventions with the intellectual disability population are often discordant with the wider perspectives of work with sexual abusers and that these men are often discriminated against.

The majority of the treatment programmes available for this group of offenders have been adapted from treatment approaches that have been successfully used on mainstream populations, in particular cognitive behavioural therapy sex offender programmes that have incorporated the specific needs of intellectual disability offenders into such interventions (Wilcox, 2004; Hayes, 1991).

Explanations of sexual offending behaviour in men with intellectual disabilities have been encapsulated into a number of different hypotheses including mental illness, counterfeit deviancy, impulsivity, sexual abuse and tendencies towards sexual offending and lack of discrimination (Craig, Stringer & Moss, 2006). For instance, the *mental illness* hypothesis assumes that these men have a higher likelihood of a dual diagnosis of mental illness which acts as a disinhibitor to offending (Craig, Stringer & Moss, 2006). The *counterfeit deviancy* hypothesis infers that sexual deviance is a result of factors such as a lack of sexual knowledge, poor social and interpersonal skills and limited opportunity to establish sexual relationships (Craig, Stringer & Moss, 2006). The *impulsivity* hypothesis is based on the assumption that intellectually disabled offenders are more impulsive than non-intellectually disabled offenders, whereas the *sexual abuse* hypothesis assumes an

association with sexual abuse in childhood and sexual offending, as a higher number of intellectually disabled sex offenders have reported having experienced childhood sexual abuse (Lindsay et al., 2001). On the other hand the *tendencies towards sexual offending and lack of discrimination* hypothesis, suggests that persistent sexual offending results from deviant sexual interest, mediated by distortions and selective cues (Craig, Stringer & Moss, 2006).

Sex offender treatment effectiveness

Wilcox (2004) suggests that treatment approaches for intellectually disabled sex offenders need to be modified to take into account poor abstract reasoning abilities and insight, whilst incorporating core elements of accredited treatment programmes, such as cognitive distortions, offence cycle and relapse prevention. Self-control skills, interpersonal skills, sex education and cognitive restructuring have also been argued to be necessary components that need to be integrated into treatment programmes for intellectually disabled sexual offenders (Coleman & Haaven, 2001).

Sex offender treatments specifically designed for intellectually disabled offenders have received some support with desirable post-treatment changes. Post intervention follow ups have seen reductions in pro-offending attitudes and beliefs (Lindsay, Marshall, Neilson, Quinn & Smith, 1998; Lindsay and Smith, 1998; Lindsay, Neilson, Morrison & Smith, 1998; Lindsay, Olley, Baillie & Smith, 1999; Keeling, Rose & Beech, 2006), a reduction in denial and minimisations, a movement towards a more external locus of control (O'Connor, 1996; Rose et al., 2002) and significant improvements in socialisation skills (Craig, Stringer & Moss, 2006). This suggests that individuals with intellectual disabilities do have the necessary skills and ability to be able to address the cognitive aspects of sexual offending behaviour. However, some studies have found no changes in the levels of cognitive distortions (Craig, Stringer & Moss, 2006) or victim empathy after participating in treatment groups (Keeling, Rose & Beech, 2006). Similarly, attitudes towards blame were found to be most resistant to change (Lindsay, Neilson, Morrison & Smith, 1998).

There is however, currently no evidence to definitely favour one treatment approach over another for sexual offenders with intellectual disabilities. Interestingly, Ashman and Duggan (2003) conducted a systematic review of interventions for intellectually disabled sex offenders and argued that the efficacy of modified programmes for sex offenders with intellectual disabilities is unclear, due to a lack of

randomised control trails of interventions. Given that it is unethical to withhold treatment, it is therefore difficult to provide controlled comparisons (Lindsay, Neilson, Morrison & Smith, 1998). Few treatment studies have been empirically evaluated, so it is unclear as to which treatment components are the most effective in reducing sexual recidivism for those with intellectual disabilities (Craig, Stringer & Moss, 2006).

Evaluating research into the treatment of sexual offenders with intellectual disabilities is further thwarted by inconsistencies in the definition of intellectual disability and sexual abuse (Courtney & Rose, 2004). There are only a few treatment outcome studies and these have been based on very small sample sizes, making generalisations difficult (Courtney & Rose, 2004). Research has also been undermined by short follow up periods and a lack of controlled treatment design (Courtney & Rose, 2004; Lindsay, Neilson et al., 1998; Lindsay, 2004). Furthermore, studies measuring treatment change have included psychometric assessments that have failed to be empirically validated on intellectually disabled sex offenders.

Social skills training

Social skills deficits have been linked to general offending behaviour (Hollin & Trower, 1986), juvenile offending (Henderson & Hollin, 1986) and sexual offending behaviour (Marshall, 2001). Insufficient social skills have been regarded as pivotal in the propensity to engage in sexually offensive behaviours, as sexual offenders lack the capacity to meet their needs in a pro-social way (Burgess, Jewitt, Sandham & Hudson, 1980; McFall, 1990). In a study by Seidman, Marshall, Hudson and Robertson (1994) different types of sexual offenders (child molesters, rapists and exhibitionists) were compared on measures of intimacy and loneliness. It was found that all groups of sexual offenders within their sample tended to exhibit high levels of loneliness and scored low of intimacy levels. Similarly, child molesters and rapists were both found to be unassertive (Segal & Marshall, 1985) and poor at reading social cues from other people (Segal & Marshall, 1986). Furthermore, social anxiety and under-assertiveness was also found in a group of child molesters (Marshall, Barbaree & Fernandez (1995). However, Overholser and Beck (1986) failed to confirm any significant differences in levels of assertiveness in sexual offenders compared to non-sex offenders and non-offenders.

Social skills are of particular relevance for the intellectually disabled sex offender population as limited relationship skills, high levels of loneliness and low self-esteem (Boer, Gauthier, Watson, Dorward & Kolton, 1995), and poor social functioning (Griffiths, Hindsburger, & Christian, 1985) were argued to be more significant for these men in comparison to mainstream sexual offenders. This is in accordance with the counterfeit deviancy hypothesis which emphasised poor social and interpersonal skills as possible explanations for sexual offending in men with learning disabilities.

Day (1994) advised that treatment and management of offenders with intellectual disabilities needed to reflect on issues of socialisation, lack of social skills, poor self-image, and poor internal controls. Similarly, Coleman and Haaven (2001) identified interpersonal skills as one of the necessary treatment components that needed to be integrated into treatment programmes for intellectually disabled sexual offenders. Haaven and Coleman (2000) also noted that the treatment strategies for sexually inappropriate behaviour typically involve social skills training.

Mulloy and Marshall (1999) noted that the early specific interventions to enhance social competence in sex offenders showed improvements in social skills. However, whilst such improvements were maintained at a two year follow-up, there was little evidence to suggest that such social skills programmes impacted on subsequent recidivism (Crawford, 1981). More recently, research has found some support for the effectiveness of a treatment component for intimacy deficits. Intimacy was significantly enhanced and loneliness reduced in a group of child molesters post treatment (Marshall, Bryce, Hudson, Ward & Moth, 1996).

The multi-component nature of treatment programmes for sexual offenders often includes some elements of social skills training, however the effectiveness of this component is rarely evaluated in isolation. Instead, treatment effectiveness is reported in terms of changes in cognitive distortions and victim empathy, with recidivism rates reported as an outcome measure for treatment efficacy. Marshall (2001) also criticised sex offender treatment programmes for being vague about the types of social skills problems sexual offenders have, despite having a component that explicitly addresses these issues, subsequently making programmes hard to replicate when the target deficits are not clearly specified.

The literature on social skills training for non-offending individuals with intellectual disabilities has shown that an assertiveness training programme produced

improvements in assertive behaviours and changes in locus of control (Bregman, 1984). Whilst studies have been successful in changing target behaviours, Robertson, Richardson and Youngson (1984) argued that very few studies reflected on other aspects of social skills such as cognition, motivation, and perception, or sought to modify the environment in order to provide natural incentives for appropriate social behaviour, given the importance of an individual's social environment. Research was further criticised for focusing exclusively at the 'performance' level of social skills training, for example on isolated pieces of behaviour such as eye contact or hand waving (Robertson, Richardson & Youngson, 1984). As an intervention strategy social skills training has produced moderate effect sizes in meta-analytic literature for those with disabilities (Gresham, 1998). Group social skills training programmes have been criticised for disregarding the type of social skills deficits the individual has, with a "one size fits all mentality". Gresham (1998) further noted that social skills training also ignores the social context within which social skills occur and lack of generalisation in a naturalistic setting.

This case study aimed to design a social skills treatment strategy specifically for a sexual offender with intellectual disabilities, in order to improve his social and interpersonal skills and subsequently reduce his stable dynamic risk level.

General Information

1. Client details

Mr J. a 48 year old single male was convicted of indecent assault against a 15 year old female. The victim was unknown to him and was on a work experience placement at the hotel where Mr J. worked, when he touched his victim on her bottom. He received a two year Probation Order, with the condition of attending a sex offender treatment programme.

Mr J. attended the Adapted Sex Offender Treatment Programme (ASOTP) at West Midlands Probation, for intellectually and socially low functioning offenders. The ASOTP is presently an unaccredited programme, as it is being piloted in a community setting. It was originally used within the HM Prison Service and was developed and adapted based on the Core Programme, for the inclusion of men not eligible for the mainstream group, which requires a large written component and the use of complex spoken language. The Core Programme is aimed at Medium to High-

risk sexual offenders with an IQ of 80 and above. The ASOTP is manualised to ensure standardised treatment delivery, and was designed to target belief systems that may be associated with sexual offending and help offenders develop a greater awareness of victim harm, to increase sexual knowledge, target cognitive distortions and develop relapse prevention skills. The ASOTP which was first introduced to the Probation Service in January 2006, consists of four blocks which are comprised of ten sessions, and each block is repeated four times, and so treatment consists of a total of 40 sessions. Some of the sessions are entitled 'making excuses', 'my four stages', 'my offence', 'new me thinking', 'risky things' etc. None of the sessions focus exclusively on social skills enhancement and probation staff were only given two weeks training on programme delivery.

Since admission onto the ASOTP, treatment facilitators deemed Mr J.'s progress as minimal and raised concerns that he had come to the completion of the programme after one year of participation and had not been able to retain any of the information discussed in previous sessions. A referral was therefore made for an assessment and for additional individual treatment, in order to reduce his level of risk.

ASOTP treatment progress reports cited emotional isolation as a continual on-going risk factor for Mr J. and that the key areas needed to be addressed were improving his relationship skills and socio-affective functioning, by using his spare time constructively.

1.1 Family history

Mr J. reported that his mother had been employed as a cleaner in a public house and his father was a watchman in a loading bay. Both his parents are now deceased. His father died in 1974 from liver problems and his mother died of natural causes in 1997. Mr J. reported being particularly close to his mother compared to his father, whom he described as 'a bit strict'. The only problem Mr J. reported with his parents during his childhood was when his father would drink heavily and argue with his mother. Mr J. described how he would attempt to defend his mum during these arguments, which turned to violence on occasions.

Mr J. has one younger brother, now aged 45, with whom he has no contact since the conviction of his index offence. He currently has no family support. Mr J. recalled that after his father died his brother became violent towards his mother and punched her, resulting in a caution from the police.

1.2 Relationship history

Mr J. was aged 21 when he had his first non-sexual relationship, which lasted only a couple of weeks. Mr J. stated that he had terminated the relationship after she continually discussed the idea of getting engaged. Mr J. entered into his next relationship at age 31. He reported that they were together for 14 years and that during this time they did not live together, although he had stayed over at her house on one occasion and had slept on a sofa. Mr J. explained that during this time they had sexual intercourse once, attributing this to his girlfriend having had a hysterectomy. Mr J. provided differing accounts as to the reason that the relationship broke down. Mr J. initially maintained that he ended the relationship, although later stated that she had been the one to end the relationship. He stated that she had owed him £1000, which he had given to her for flowers and jewellery and when he asked for it back they had gotten into an argument. He added that he had wanted to get engaged to her but she had refused because of being negatively treated in previous relationships.

1.3 Educational and employment history

Educationally, Mr J. recalled that he had attended a mainstream school for the first three years, but was subsequently moved to an 'open air school' for children with learning difficulties due to his epilepsy. Mr J. stated that in 1967 he had an ear inflection which triggered epilepsy. He recalled that he had seizures once a week and had been given 'valium and another drug' to control his epilepsy. He described how the medication had negative side-effects that would cause him throw objects about.

Throughout his school years Mr J. was continually subjected to teasing and bullying on a weekly basis. This involved threats of physical harm if he did not pay his bullies with his dinner money, yet Mr J. still considered them to be his 'bad' friends. Upon leaving school with no formal qualifications Mr J. maintained a relatively stable employment history. He first worked in a warehouse packing fruit for four years, until he got the sack after being caught giving boxes of bananas to children. The following five years Mr J. remained unemployed and received benefits until he managed to find work as a landscape gardener. This job lasted a year until his contract come to an end.

Mr J. then found employment as a kitchen porter in a hotel, where he worked for over 17 years, until the hotel lease came up for renewal. During this period he recalled that his line manager at work bullied him for money, and he was forced to go

to a cash machine and withdraw money whilst his line manager stood behind him. Mr J. stated that he felt angry and wanted to 'kill him', but gave him the money because he was frightened of him. In 1999 Mr J. found a similar job as a kitchen porter in another hotel and worked there until he was suspended for committing his index offence. He has since remained unemployed.

1.4 Medical and psychiatric history

Mr J.'s medical history indicated that he has suffered from epilepsy during his childhood, however, his seizures ceased during his early adulthood. He has no history with the psychiatric services.

1.5 Forensic history

Mr J. was first convicted of indecent assault at the age of twenty one. He was held on remand and made subject to a two year Probation Order. His victim was a seven year old girl. He states that he touched her in her private parts and that the police had also found twenty one pornographic books. Shortly after this incident Mr J. moved to Birmingham.

Mr J. further disclosed that he had been caught by a security guard whilst exposing himself in bushes in a park, and taken to the police station. He was unable to recall the exact year but stated that it was not long after moving to Birmingham and he had been 'let off' for that offence. Mr J. did not receive another conviction until his index offence of indecent assault. Mr J. has no other criminal convictions and does not hold antisocial cognitions that support general criminal behaviour.

2. Assessment

This case study method was based on the ABA design. An initial pre-intervention baseline assessment was completed. This has been used to construct a formulation of the presenting problems, directing the course of the intervention. Assessments were then re-administered to evaluate the effectiveness of the intervention.

2.1 Archival sources

An in-depth file search was carried out of all available documents. File based information contained treatment notes and Pre-Sentence Report and information regarding his offence.

2.2 Interview

A semi-structured interview had been completed, including family, forensic and educational histories. Details gained from the interview and archival sources were used to gain collaborative information, in an attempt to gain an understanding of Mr J.'s current difficulties. Adaptive functioning was assessed during the interview process, as psychometric tests measuring adaptive functioning were not available. This information can be seen in both the client details section and the following functional analysis. The purpose of the evaluation was reiterated throughout the assessment so that the client did not misinterpret the purpose.

Behavioural observations were limited to the time spent during the assessment as Mr. J was not in a residential setting where detailed observations could be conducted. The interview setting was at West Midlands Probation Services and was conducted in a small private interview room. During the interview cognitive and memory deficits were apparent.

2.3 Psychometric tests

Psychometric assessments will be discussed in terms of the assessment of needs and responsivity. Responsivity was examined through the assessment of intellectual functioning and adaptive functioning. The assessment of need included an assessment of dynamic risk (both acute and stable). Acute factors fluctuate and change, signalling the onset of offending, for instance mood state and substance abuse (Hanson & Harris, 2000; 2002). Stable factors on the other hand are those that may be amenable to change in treatment such as pro-offending attitudes, deviant sexual interest, and socio-affective problems (Thornton, 2002). By addressing dynamic risk during treatment, a reduction in sexual recidivism risk is likely to follow (Ford & Beech, 2004; Hanson & Harris, 2002).

A battery of self-report adapted psychometric tests were administered in order to help aid in the prediction of dynamic risk and dangerousness. Psychometric assessment holds the advantage over other forms of assessment, as information can be obtained in a standardised and valid way within a short amount of time. Questions within the psychometric tests were presented verbally due to the reading difficulties of the client, an appropriate method of assessment identified for this population (Boer et al., 1995).

For the purpose of this case study only the psychometrics indicative of a problem will be included and discussed. It should be noted that other psychometric tests were administered but excluded due to the client's lack of comprehension (for example, when asked to clarify his answers, it was apparent that he had not understood the questions, and so the tests were not valid).

2.3.1 The assessment of need

Paulhus Deception Scales (PDS; Paulhus, 1984)

This is a 40-item questionnaire, which had recently been included for the adapted assessment battery by HM Prison Service, for sexual offenders with intellectual disabilities. It measures the tendency to give socially desirable responses, and containing two subscales; Self-Deception and Impression Management. The *Self-deception* scale provides information on the tendency of some respondents to provide agreeable self-profiles that are due to an overly confident self-regard and the *Impression Management* scale provides information on the tendency of some respondents to consciously respond to items in an attempt to make themselves appear favourably to those interpreting their results. The questionnaire is scored using a 7 point likert scale ranging from 1 (not true), 4 (somewhat true) to 7 (very true), representing how much the statements are true of them.

It is important to note that psychometric properties of the scales effectiveness have yet to be established for specific use with intellectually disabled offenders. However, psychometric characteristics of the PDS based on a sample of prison entrants, indicates the instrument's internal reliability appears to be high, ranging between .72 and .84 (Cronbach's alpha) for both scales (Paulhus, 1984).

Mr J. obtained a score of 10 on the Impression Management scale, although normative data for an intellectually disabled population has yet to be published, a score of 10 on this scale is within the fake good range, indicating that Mr J. is attempting to present himself in a positive light. On the Self-Deception scale he scored 5, which is within normal limits.

Rathus Assertiveness Schedule (RAS; Rathus, 1973)

The Rathus Assertiveness Schedule is one of the most well-known self-report measures of assertion (Bellack & Hersen, 1998). This is a 30-item questionnaire rated on a 6 point likert scale (see Appendix 12), ranging from +3 (very characteristic of

me) to -3 (very uncharacteristic of me). Some of the items have reversed scoring, with higher scores reflecting more assertiveness. The total score is the sum of all the items, yielding a total assertiveness score of between -90 and + 90, with positive scores indicates assertiveness and a negative score indicates non-assertiveness. Rathus (1973) reported moderate to high test-retest reliability (r =.78, p<.01). Normative data based on 68 undergraduate students indicated a mean pre-test score of-0.3 (SD = 29.1). Overholser and Beck (1986) compared sexual offenders on the RAS, and found that child molesters had a mean score of -10.3 (SD = 20.9) compared to a sample of college students who had a mean score of -2.1 (SD = 24.5).

Mr J. obtained a score of -19, which is below average and indicates that he lacks assertiveness. This score is very low in comparison the general population and the child molesters in Overholser and Beck (1986) study. Visual inspection of the items indicated that Mr J. had problems with refusing requests and expressing his opinion. The RAS has no normative data for offenders with intellectual disabilities.

Short Self-Esteem Scale (SSES; Thornton, 1995)

This is an 8-item measure evaluating levels of self-esteem, using a dichotomous yes/no scale. A 'yes' response was scored as 2 and a 'no' response was scored as 0 (see Appendix 12). It is a useful screening tool for self-esteem deficits in sexual offenders. The questionnaires has test-retest reliability score of *r* =.75 (SD = 2.71) as reported by Beech, Fisher and Beckett (1998). More recently Webster, Mann, Thornton and Wakeling (2007) found the SSES had excellent internal consistency and test-retest reliability, with a Cronbach alpha coefficient of .84 and a test-retest reliability score of *r* = .90. The mean pre-treatment score for a sample of extra-familial offenders with female victims was 7.9 (SD = 5.2), compared to a non-offender sample who produced a mean score of 15.5 (SD = 1.0).

Mr J. scored 5, suggesting below average levels of self-esteem, based on a profile of non-offending adult males. Psychometric information was not available for offenders with intellectual disabilities for appropriate comparison.

2.3.2 The assessment of responsivity

The Wechsler Adult Intelligence Scale-III (WAIS-III; Wechsler, 1997) was used to assess intellectual functioning. The WAIS-III has been recommended for assessing intelligence in this population (Keeling, Beech & Rose, 2007). The WAIS is a general

test of cognitive ability and consisting of 14 subscales, measuring verbal and nonverbal reasoning abilities. The subscales scores are then combined to produce an overall estimate of general functioning. Three main scores are provided: the Verbal IQ (VIQ), a measure of acquired knowledge; Performance IQ (PIQ) a measure of fluid reasoning, spatial processing and visual integration; and Full IQ score, which is an aggregate of the PIQ and VIQ.

An intellectual disability is defined by the American Psychiatric Association (2000) as a full scale IQ score of approximately 70 or below, with the onset before the age of 18 years and concurrent deficits in present adaptive functioning. Table 4.1 shows that for Mr J.'s WAIS-III scores, his VIQ score was at the 1^{ST} percentile, falling in the extremely low range, whereas his PIQ fell within the classification of low average. Mr J.'s Full scale IQ score placed him at the borderline range of intellectual functioning, at the 5^{th} percentile rank.

Table 4.1. Summary of the WAIS-III scores

	Score Obtained	95% Confidence Interval	Percentile Rank Range	Intellectual Disability Range
Verbal IQ	67	63 - 73	.01 - 2	Extremely Low
Performance IQ	89	83 - 96	9 - 23	Low Average
Full Scale IQ	75	71 - 80	2 - 8	Borderline

In accordance with the Wechsler Intelligence Scale manual, a 20 point discrepancy between Mr J.'s VIQ and PIQ was considered statistically significant ($p<.05$), and clinically abnormal. It was possible to infer from the difference that Mr J. performance abilities exceed his verbal abilities. Mr J. did not have a hearing impairment which could have accounted for his low VIQ. When PIQ exceeds VIQ it may be suggestive of lower education, reduced attention span, problems with verbal reasoning or better nonverbal intellectual abilities. This information will be considered when determining appropriate intervention strategies.

2.4 Risk assessment

Sexual Violence Risk-20 (SVR-20; Boer et al., 1997)

The SVR-20 adopts a guided clinical assessment format to assess the dynamic risk of sexual violence recidivism. It contains 20 items divided into three risk categories: psychosocial adjustment, sexual offences, and future planning. Static, stable and acute dynamic factors are included in the SVR-20. Mr J. was judged to be at Medium risk of sexual violence, on the basis on employment and relationship problems, the loss of a social support network, and the minimisation of his sexual offence.

RM2000 (Thornton et al., 2003)

The RM2000 is a two-dimensional risk assessment for sex offenders. The first scale is the Risk Matrix 2000/Sexual (RMS), measuring the risk of sexual recidivism, and the second scale the Risk Matrix 2000/Violent (RMV) measures the risk of non-sexual violence. On the RMS Mr J. was assessed to be in the Medium risk band for committing a further sexual offence on the basis of his offence, his criminal history and other demographic factors. Reconviction rates for medium risk individuals being reconvicted of a sexual offence over a 5, 10 and 15 year period are 13%, 16% and 19% respectively (Thornton et al., 2003). Mr J. self-reported exposing himself in a park (but was not convicted) however, this would serve to increase his risk to the high risk band. He presents a Low risk for violent reoffending.

Rapid Risk Assessment for Sexual Offence Recidivism (RRASOR, Hanson, 1997)

The RRASOR is an actuarial tool containing four items. These items are past sexual offences, age at commencement of risk, extra-familial victims, and male victims. Mr J. was allocated points according to the presence of these. He obtained a score of 2, out of a possible score of 6. The higher the value of the risk scale was associated with an increase in the sexual offense recidivism rate. Hanson (1997) reported the sexual offence recidivism rate for offenders with a score of 2 was 14.2% at 5 years and 21.1% at 10 years, which was indicative of low to moderate risk.

Assessment of Risk Manageability for Intellectually Disabled Individuals who Offend (ARMIDILO; Boer, Tough & Haaven, 2004)

The ARMIDILO incorporates 30 items, which includes *stable* and *acute* dynamic factors. The stable dynamic client factors relevant to Mr J. were his poor knowledge

of his behaviour cycle, risk factors, sexual knowledge, as well as poor time management skills, general coping ability and relationship skills. Acute dynamic client factors which served to increase Mr J.'s level of risk were changes in emotional state, victim access, and changes in the ability to use coping strategies. An acute dynamic environmental factor identified was changes in social relationships.

Functional Analysis

Functional analysis is an approach to help understand the causes of problem behaviour by establishing reinforcement contingencies and operations for problem behaviour (Holden, 2002). Functional analysis examines the functional relationship between the explicit environmental and historical variables that control and affect a person's behaviour (Hopko & Hopko, 1999). Functional analysis is not a unitary method; it covers indirect measures based on observations and interviews and experimental manipulations, and serves to increase the efficiency of treatment design (Holden, 2002). However, functional analysis has limitations. Reinforcing variables may be hard to withdraw, the primary functions of a behaviour can be difficult to detect, and the behaviour can be multiply controlled (Holden, 2002).

The assessment findings have been used to formulate a functional analysis and develop a clinical hypothesis about the client's problem behaviour and used to derive a related treatment strategy. The following section summarises what appears to be the causes of Mr J.'s current social skills problem. Figure 2 (page 82) indicates that Mr J. witnessed parental conflict and verbal and physical abuse towards his mother. His father was often unavailable due to his daily drinking. These experiences resulted in Mr J. suffering from low esteem as he felt inadequate and unable to protect his mother from the violence.

Mr J. suffered from epilepsy during his childhood which may have negatively impacted upon his cognitive development, as brain damage may have occurred due to the lack of oxygen or possible head trauma during his seizures. He was subsequently sent to a 'special school' due to his intellectual difficulties. Whilst at school he was systematically bullied and rejected by other pupils, resulting in social isolation which impacted upon his self-esteem. Due to social isolation whilst at school he lacked the opportunities to engage in and practice interpersonal skills. Mr J's poor interpersonal and communication skills led him to receiving negative reactions from those around

him, resulting in low self-esteem and low mood. Mr J. often felt rejected and developed dysfunctional assumptions about himself. Inadequacies in peer relationships can lead to a decrease in social support, resulting in the decreased likelihood that relationships will be used as a buffer against stress (Sroufe & Rutter, 1984).

Mr J. experienced difficulties relating to others and found it particularly difficult talking to members of the opposite sex, and so was unable to establish intimate relationships during his adolescence, which subsequently lead Mr J. to seek sexual gratification form of pornography. Mr J. acquired a large collection of pornography purchased from a local news agents, although the pornography he viewed did not have a specific theme of sex and violence. However, such exposure to pornography may have resulted in Mr J. learning inappropriate and distorted messages regarding sexual behaviour and the objectification of women in the images he viewed.

Mr J. also had difficulties refusing the requests of others. An inability to assert himself accordingly, coupled with mixed feelings of fear of rejection and a desire to be accepted, led Mr J. to give others his money. This pattern of behaviour continued into adulthood, as he continued to be financially exploited throughout his adulthood. He would give part of his wages to his supervisor out to fear of physical reprisals. He would also give money to his "girlfriend", out of fear of rejection in the relationship. His lack of assertiveness was also evident in a poor awareness of his body language (slumped posture and poor eye contact). Behaving unassertively allowed Mr J. to avoid anxiety provoking situations and confrontation, and the possibility of rejection from others. Such a strategy unfortunately resulted with him feeling that he had been taken advantage of. Mr J. lacked sexual experience, he has had only one "relationship" as an adult, which he maintains lasted for fourteen years, however during this period of time he had had sexual intercourse on only one occasion.

With regards to Mr J.'s index offence (see Figure 3), it was felt that a lack of understanding of personal boundaries and space, intimacy deficits and a negative emotional state prior to committing his offence were predisposing factors. In the build up to his index offence, he experienced a change in a social relationship, as a few days prior his relationship with his "girlfriend" had ended. Another acute dynamic risk factor for Mr J. was his negative affective/emotional state. He described being in a 'daze' and that he 'did not feel good'. He had been staying over at his supervisor's house and had drunk an excessive amount of alcohol the evening before his offence.

Whilst at work the next morning Mr J.'s supervisor had asked him to do additional duties upon his normal work load. Mr J's lack of communication and assertion skills resulted in him complying with the instructions, as he felt unable to ask for extra help. Subsequently he became angry and stressed at his perceived lack of help from work colleagues, and was unable to utilise effective coping strategies. In addition to feeling 'hung-over', these proximal factors contributed to his offence.

With regards to Mr J.'s offence, he exhibited poor impulse control when he indecently assaulted his victim. He stated that he was unaware that he had committed a sexual offence until he was requested to report to his supervisor later that day. Mr J. initially admitted to having fantasies about his victim being shy around him because she liked him sexually. However, he later minimised responsibility for his offence, insisting that he was simply trying to get past his victim, and only touched her in order to move her out of the way by touching her from her waist down to her bottom. Mr J. maintained that the only thing he did wrong was failing to say 'excuse me'.

Figure 2. Functional analysis of Mr J.'s adolescent years

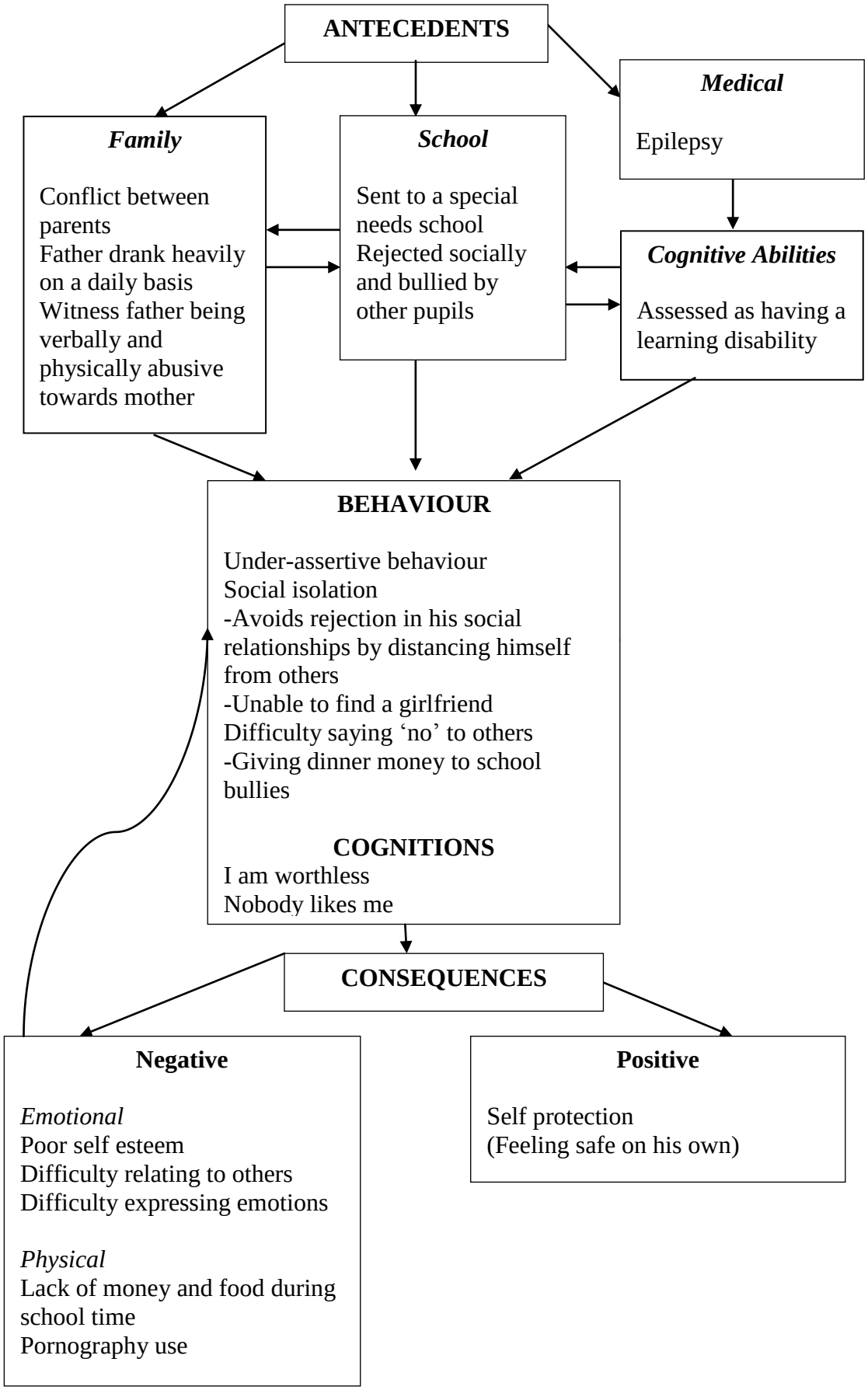
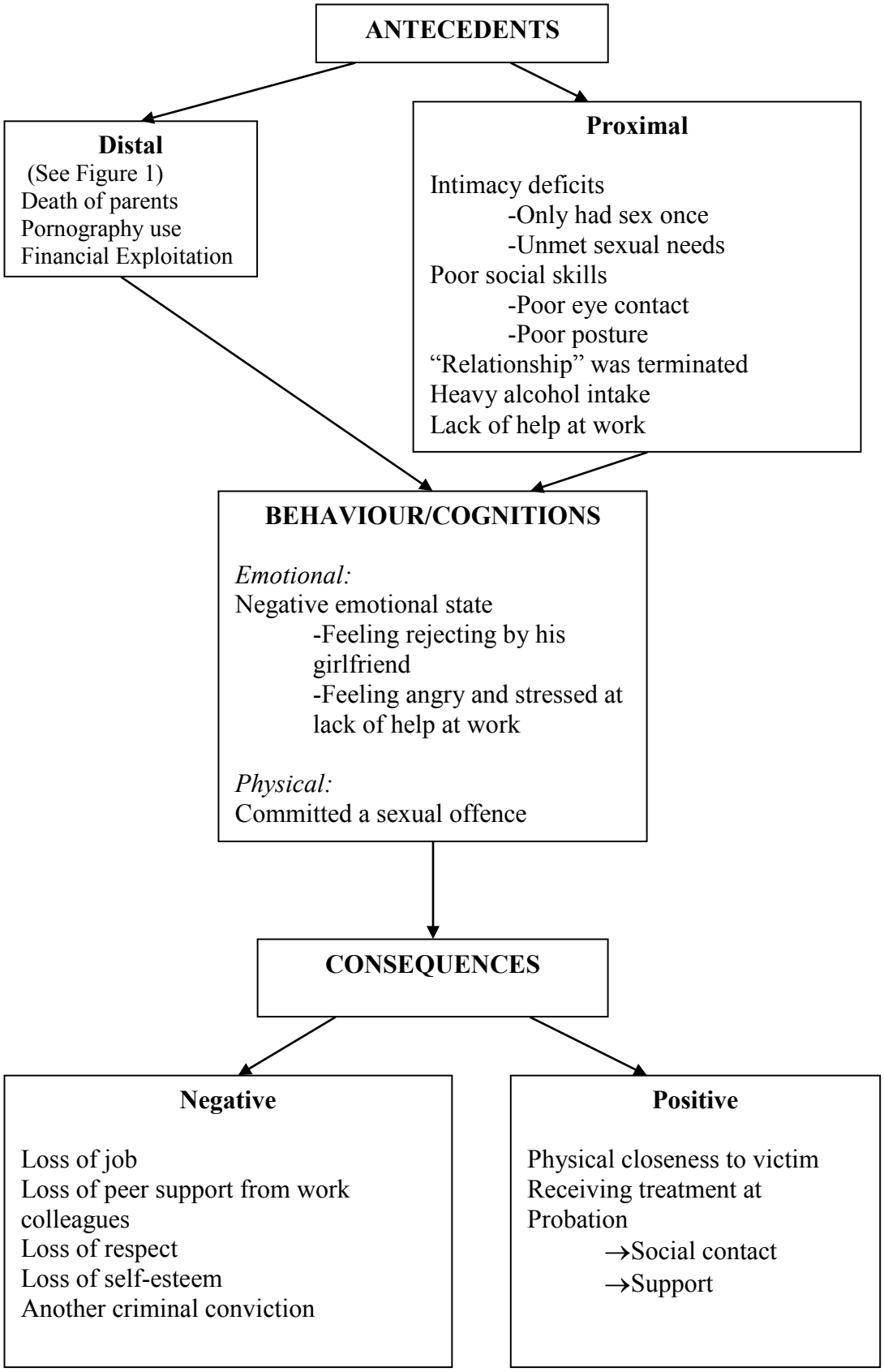


Figure 3. Functional analysis of Mr J.'s index offence



Intervention

The assessment and functional analysis indicated that Mr J. would benefit from an intervention to address social skills deficits. The intervention was designed specifically to increase the client's assertiveness, anger control, self-esteem and interpersonal skills. Each specific topic was divided into short intervals and repeated throughout, to sustain his attention. The intervention was scheduled for a regular time twice a week and each session lasted two hours, including a short break. The total number of sessions was limited to ten sessions because the client's Probation Order was due to expire shortly after the referral for individual work was made. I delivered the intervention, working under the supervision of a clinical and forensic psychologist.

Treatment strategies that encourage motivation, foster self-reliance, and create facilitating environments all increase the efficacy of treatment (Coleman & Haaven, 2001). A multi-sensory and engaging approach to treatment programmes has been shown to facilitate motivation and the ability to retain information and strategies to avoid relapse (Wilcox, 2004). The simplification of treatment concepts and the use of visual imagery are also imperative in the treatment of the intellectually disabled population (Lambrick & Glaser, 2004).

The intervention design, therefore, incorporated the use of simple language and pictorial aids where possible. A repetition and a re-cap of the previous session in the next session was done in order to aid Mr J's memory. Repetition of the key points and behaviours was considered important, as the literature suggests that in the presentation of treatment, the use of constant repetition, limited amounts of information and the use of concrete examples should be used in the treatment of those with intellectual disabilities (Clare, 1993). Cognitive and motivational aspects of social skills were also incorporated into the treatment strategy. Verbal praise was given to the client when he had an understanding of concepts and when he demonstrated positive behaviour.

The client was introduced to a diary at the beginning of the intervention and was asked to record social interactions, rating the duration, quality and effectiveness of the interactions. Reminder cards, role-plays, feelings chart, 'old-me', 'new me' concepts, and drawings were some of the techniques used throughout the intervention.

Treatment sessions

Session 1: Introduction session

The client was helped to explore body language, verbal and nonverbal behaviours (such as physical posture, tone of voice and facial expressions). Personal space was discussed with the client, using the idea of an invisible bubble which expands and shrinks depending on the relationship with the person and the situation. Personal boundaries were discussed in relation to inappropriate touching (when and with whom is it appropriate to touch a person). Pictorial aids were used and the client was asked to mark in colour on a picture of a man and women locating where on the body it was acceptable and not suitable to touch.

Mr J. was encouraged to make more eye contact during the sessions as his eye contact was extremely poor during the assessment. Every time he was successful at making eye contact it was recorded as an eye lash on a drawing of an eye, with the aim of eventually producing an eye full of eyelashes in a session. It was hoped that this visual imagery would be a fun and constant visual reminder of the importance of making eye contact, as well as a record for him to see his own progress (see appendix 14 for eye contact sheet and worksheet materials).

Session 2: Non-verbal behaviour

The work from the last session was discussed, repeating the importance of personal space and body language. Mr J. stated that after the last session he made conversation with an elderly lady on the bus home and said that he made efforts to maintain better eye contact. The client was verbally praised when making appropriate eye contact.

In this session three behaviour types were introduced: passive, aggressive and assertive. The detrimental effects of passive and aggressive styles were explored. The client practiced assertive responses and learnt to distinguish assertive responses from compliance and aggression. Again the importance of body language in social encounters was discussed with the client in relation to assertive behaviour. Diagrams were used to show important areas such as eye contact, posture and hand gestures. Body language was explored in relation to assertiveness. The client was encouraged to maintain eye contact and this was again recorded. The client was verbally praised whenever he made efforts to make eye contact.

Throughout the session the client was encouraged to practice his eye contact and posture. The client was asked to identify his hobbies and was encouraged to engage in more social activities by joining a club so that he could practice his social skills with peers. Leisure-time skills were considered important for Mr J., as boredom and poor social skills were viewed as likely risk factors for committing another sexual offence. Hobbies and activities would aid directly in relapse prevention work, as well as help manage his time more effectively.

Mr J.'s homework assignment was to practise assertiveness skills and to record his social interactions in a diary provided (see Appendix 15). He was also set the goal of joining or planning to go to a snooker club which an acquaintance attends.

Session 3: Learning to say no

The work from the previous session was discussed. He had completed his diary from the previous session and the social interactions he recorded were evaluated. The aim of this session was to provide the client with skills to enable him to respond assertively to unwanted requests. Body language, not making excuses and not apologising in relation to refusing a request were explained in relation to acting assertively. The 'broken record' technique was also explained. The client was helped to identify situations in which he has difficulty saying 'no', such as market researchers and bank personnel getting him to sign up for credit cards (he had acquired eight cards in total). The consequences of not saying 'no' to somebody were examined, i.e., feeling angry.

Immediate rehearsal after learning was done via role-plays of these situations, whereby appropriate behaviours and responses were modelled. The 'old me' and 'new me' technique was incorporated into the role-play situations. We role-played how he previously responded and then rehearsed how he could assertively respond in the future by practicing saying 'no'. Mr J. was rated on three areas, his eye contact, posture and his ability to refuse requests, and feedback was provided. Mr J. performed well at saying no using the broken record technique but had difficulty maintaining eye contact and good posture during the role-plays.

Session 4: Dealing with bullying

The content of the previous session was discussed, repeating the importance of eye contact, posture and saying 'no' to others. Diary entries of social interactions were

recorded. Mr J. was asked to discuss how successful he had been and what he could have done differently. Mr J. identified that he had not made "full eye contact" because he had kept his head down whilst talking to an old work colleague. He felt he performed poorly because he was surrounded in a busy crowd and could not concentrate.

During the session the client had brought with him an eight page list of leisure clubs, which he had obtained from staff at a public library. Time was spent time discussing the list of clubs that he wanted to join, the potential risks in terms of children also attending were also discussed. The client was praised for taking the initiative and such a proactive approach. He was encouraged to join an adult only groups, and appeared very enthusiastic about joining a social club.

The aim of the session was to increase the client's awareness of bullying and its effects on him and to help him. Mr J. identified how other people had been 'nasty' to him in the past and discussed other types of bullying behaviour. Mr J. was then asked to draw the emotions he felt onto a picture of a blank face. He also drew a picture of a sad face and talked about how he used aggression to keep others away, which had made him feel lonely. How he had dealt with bullying in the past (Old Me) and what he could do differently (New Me) was explored.

Session 5: Dealing with bullying

The work from the previous session was discussed. Mr J. was unable to complete his diary since the last session as he had had no social encounters. The eye contact sheets from the previous sessions were presented to Mr J. so that he could see how his eye contact had improved, since the start of the intervention.

The session continued from the last session, examining in more detail the emotions and thoughts he has had in response to bullying i.e., 'if nobody wants me I would rather be on my own', and how this linked together and resulted in him isolating himself from others.

The idea of positive self-talk statements was introduced as a coping strategy to use when bullied. The client came up with statements he could use, such as 'I am ten times better than them', instead of thinking about being on his own, and he practiced saying these statements. Alternative responses to being bullied in the future were brainstormed with the client, for example, walking away, not answering back and

reporting it to somebody. His previous actions were discussed in terms of how this made the situation worse, and how he can stay calm in future was examined.

Session 6: Anxiety management

During this session the client was asked to explore different emotions (feeling anxious, angry, sad and happy). He was asked to draw what each face would look like when feeling these emotions and to identify how the body responds to such emotions and how he would behave. He identified feeling anxious and angry most of the time. He then identified which situations trigger these feelings. Mr J. was encouraged to identify the physical and psychological signs of anger arousal he experiences and what triggers his angry feelings.

The rest of the session focused on what he could do in response to these negative emotions. It was explained how physical exercise can help reduce anger and anxiety. Three different relaxation techniques (controlled breathing, physical relaxation by tensing muscle groups and visualising calming pictures) were introduced to Mr J. He practised each of these techniques a couple of times. Mr J. reported that he found the calming pictures particularly useful. He was asked as a homework assignment to practice the self-instructed relaxation techniques over the weekend.

Session 7: Conversation skills

The work done from the previous sessions was briefly discussed, particularly with regards to communication and assertiveness skills. During this session my practice supervisor participated in the session in order to provide an opportunity to assess Mr J.'s eye contact. Issues of personal hygiene in relation to personal presentation were emphasised, and therapeutic boundaries were discussed with the client in a sensitive manner.

The focus of the rest of the session was on communication skills. Open and closed postures were discussed and practiced, as were nodding and hand-shaking. Personal space was also explored with Mr J. The idea of 'hot buttons' as a way of initiating and maintaining conversations was introduced (hot buttons are areas of interest to create enthusiasms in people you talk to, such as work, hobbies, sporting events etc). Mr J. identified his own 'hot buttons' that he could use in conversation when meeting potential friends. Mr J. was encouraged to put into daily practice the

skills he has acquired. It was emphasised to him that sometimes his attempts may be met with failure and rejection but that he should keep trying and not to take it too personally. Mr J. was encouraged to practice and increase the frequency of his social contacts.

Session 8: Problem solving

Aspects of the previous sessions were repeated to help consolidate the information that he had learnt so far. During this session Mr J. reported feeling anxious, and the nature and extent of his problem was then discussed. He identified that his flatmate was not paying his share of the bills and as a result Mr J. was receiving fines, and this had been a reoccurring problem. The client was helped to generate solutions to the problem that were in the interests of both parties. Techniques and skills from previous sessions were incorporated and the client was encouraged to role-play assertive responses in preparation. Mr J. was encouraged to use the relaxation techniques he had learnt prior to dealing with his problem.

Session 9: Facing your fears

This session continued to explore his barriers to making friends and how to overcome this. During this session Mr J. was encouraged to looked at how his feelings of anxiety had made him avoid people and withdraw by staying at home, and how in the short term this had provided immediate relief, but in the long term it had made him feel worse because he felt lonely and wanted to do destructive things in his home. The idea of becoming more active as a coping mechanism was also discussed. Mr J. expressed that he wanted to be more socially active because of the loneliness he was experiencing.

The task of meeting people new people and making friends was the focus of the remainder of the session. This task was then broken down into eight smaller steps he needed to take in order to achieve this goal. By breaking this down into smaller steps Mr J. may find it more manageable. He acknowledged that he had made the first step by approaching the library and requesting information for information on clubs. The next step he identified was then to phone up a club and ask for information, the next small step was to find out how to get to the venue and the final step was to actually go and talk to somebody. At this stage Mr J. was encouraged to practice his relaxation exercises when he became anxious.

The formation of friendships was also addressed, from becoming a stranger, to an acquaintance and then to friendship. Personal boundaries within friendships was again emphasised along with the role of professionals and how they were different from friendships.

It was hoped that improvement in self-confidence would enhance the quality of future relationships, so at the end of the session self-esteem enhancing procedures were implemented, for example the client was asked to list 5 or 6 personal features that are positive. A goal was then established with the client, that he was to try and rehearse each of the positive self-statements at least 3 times a day over the following week.

Session 10: Overview of previous session

Mr J.'s diary sheets and eye contact records over the course of the intervention were examined and discussed. An overview of what he had learnt in the work done in each of the previous sessions was provided. Relapse prevention work was also incorporated into the final session, focusing specifically on maintaining personal boundaries and not encroaching on other people's personal space through inappropriate touching.

Summary of observations during sessions

The client commented that he thought the assessment/intervention process was useful as it made him think about himself more. The client was very responsive to the eye contact chart and appeared to enjoy it. More importantly his eye contact during the sessions increased dramatically. Mr J.'s level of eye contact, however, was still not at a normal expected level, he continued to have a tendency to avoid gaze when in conversation and when not being reminded to make eye contact. At first he needed a lot of prompts and reminders about making eye contact, although towards the end of the intervention he did not require as much. Mr J. expressed feeling pleased with his progress on making eye contact.

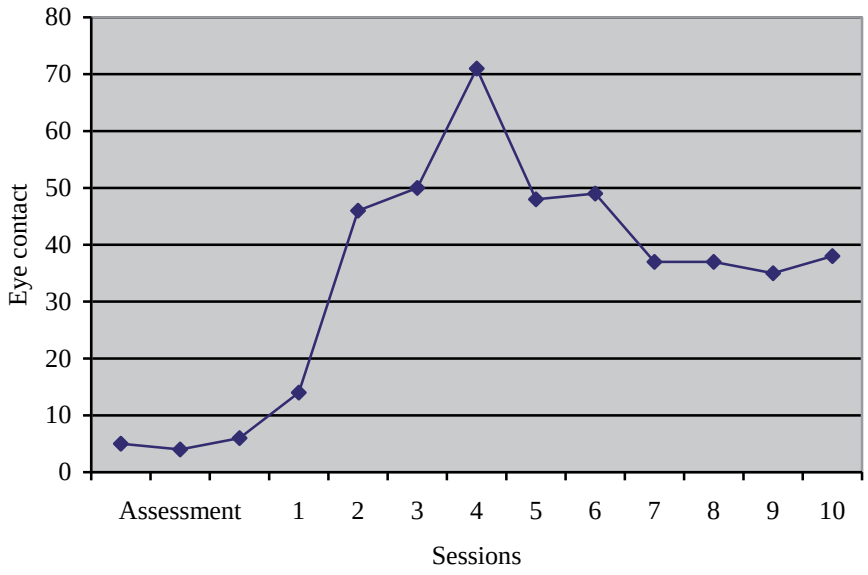
Issues of dependency were apparent during the intervention. In one of the sessions Mr J. stated that he considered me as a friend because he felt at ease talking to me. Boundaries between the therapist and the client were explored in a sensitive way and the aims of the intervention were reiterated to the client, i.e. to help him be able to make friends more easily.

Results

Eye contact evaluation

Figure 4 shows a dramatic improvement in eye contact after the assessment period when the eye contact charts were introduced. After Session number 6, it was agreed with the client that the number of reminders and prompts to make eye contact would be reduced, as in a natural social setting people would not remind him to make eye contact, at this point his level of eye contact stabilised. It should be noted that whilst this level of eye contact had subjectively improved it is still relatively poor, and whilst the frequency of eye contact was recorded unfortunately this was limited as it failed to show the duration of the gaze made.

Figure 4. Chart showing the number of times the client made eye contact



Psychometric evaluation

Treatment impact was assessed by looking at whether Mr J. had shifted in his level of assertiveness and self-esteem following treatment on the psychometric tests. Responses are examined at an individual level to ascertain whether he had moved from a score more likely to be found in a dysfunctional distribution of scores (e.g., child abuser attitudes), to a score more likely to be found in a functional distribution of responding (non-child abuser attitudes). This methodology was chosen as it is a standard method used in a number of other areas to assess the impact of therapy

(Kazdin, 2003). There are a number of systems to assess individual change (see Kendall, 1999) but the method employed in this study is the same as that described by Jacobson and his colleagues (Jacobson, Follette & Revenstorf, 1984; Jacobson & Traux, 1991). Jacobson suggests that in order to assess significant change, two aspects need to be determined: (a) whether an individual's score after treatment is within the cut-off point between normal and dysfunctional responding on the particular measure of interest; (b) whether the amount of change pre- to post-treatment was statistically reliable.

The cut-off between dysfunctional and functional responding was assessed as follows using the Jacobson and Traux (1991) system:

$$\text{cut-off} = \frac{(SD^1)(MEAN^2) + (SD^2)(MEAN^1)}{SD^1 + SD^2}$$

Where $MEAN^1$ and SD^1 are the mean and the standard deviation of functional groups (i.e., non-offenders) and $MEAN^2$ and SD^2 are the mean and the standard deviation of a dysfunctional group (i.e., child molester offenders). For the RAS the cut-off was calculated using data from Overholser and Beck (1986). The cut-off for the Self-Esteem measure was based on data from Webster et al. (2007).

The second part of this system measures whether the amount of change on a particular treatment measure was statistically reliable. The Reliability of Change Index (RCI) was employed which has been described by Jacobson et al. (1984) as:

$$RCI = \frac{\text{(post-treatment) - (pre-treatment)}}{S_E}$$

The method of calculating S_E was as follows: $S_E = SD_x \sqrt{(1 - r_{xx})}$. With r_{xx} as the test-retest reliability of the measure in question and SD_x represents the pre-treatment standard deviation for the measure in the offender sample. An RCI greater than 1.96 (standard error of measurement) represents significant pre - post change (i.e., the 5% significance level for a one-tailed test. It should also be noted that Christensen and Mendoza (1986) suggest that S_E should be replaced S_{diff} to reflect the amount of difference which one could expect between two scores, obtained on the same test by the same individual, as a function of measurement error alone. The S_{diff} was calculated using the method; $S_{diff} = SD_x \sqrt{2(SE)^2}$.

Table 4.2. Client's pre and post treatment scores

Psychometric Test	Cut-off Score	Pre-Treatment Score	Post-Treatment Score
Self-Esteem	14.4	5	6
Rathus Assertiveness Schedule	-6.5	-19	-24
Paulhus Scale (IM)	-	10	11

Table 4.2 above shows Mr J.'s test scores taken before and after completion of the treatment intervention. Re-administration of the psychometric tests indicated no significant clinical change, using the above described formula. Calculations indicated that Mr J. had not changed clinically post intervention on the Self-Esteem or RAS questionnaire. On the RAS a five point different was obtained, although this was marginal and not indicative that his assertiveness had got significantly and clinically worse. His test scores remain outside what can be defined as normal functioning and his post treatment scores were not above the cut-off score for clinical change.

It should be noted that the cut-off and the RCI could not be calculated for the Paulhus Scale, as test-retest data for this newly adapted scale was not available, although his score remained fairly consistent, indicative of no change, given that it was only a one point difference.

Discussion

This case study was a discrete piece of work aimed to increase the client's social skills and competencies. Within the time-frame of the case study only very specific aspects of the client's problem behaviour in relation to social skills could be targeted as the intervention needed to be pitched at a slow pace with repetition being a key element. However, this client has multiple treatment needs and ideally individual work would continue on a longer term on-going basis to address other important issues. It was hoped that by improving the client's social skills he may be able to go back into the adapted treatment programme and repeat a module. The client's lack of social skills was also considered to be linked to his offending behaviour.

The client achieved some improvements in social functioning, for instance his level of eye contact was markedly better. However, re-administration of the psychometric tests did not show significant improvements to his level of assertiveness. It should be noted that on the Rathus Assertiveness Schedule his post-treatment score appeared slightly worse, although this score was fairly consistent with his initial assessment and did not represent a significant change. This small change in scores could have be attributable to the client feeling more comfortable disclosing within treatment or that during the course of sessions the treatment focus may had prompted him to rethink and became even more aware of his own under-assertive behaviour as he learnt more about under-assertiveness and body language.

Possible explanations for the treatment outcome and the problems faced during the course of the case study are discussed below.

3.1 Intervention setting and characteristics

An advantage of a community setting is that it provides greater ecological validity, enabling problems to be addressed in the real-world context when they occur (Howells & Day, 2003). However, there were a number of clinical problems associated with a community-based intervention, which may have affected the treatment outcome, mainly the contact with the client. This was limited compared to in a residential setting, where contact can be daily, thus enhancing the client's trust in the therapist and aiding the therapeutic alliance. This case study was conducted within a limited time scale, and restricted to a probation setting, it was therefore difficult to observe the client practising the skills he had learnt. A lengthy follow-up period would also have been useful to see if behavioural changes in the levels of eye contact were maintained. Recidivism data would also have been a useful outcome measure to evaluate the treatment success, as the client's social skills were connected to his offending behaviour.

Coleman and Haaven (2001) highlighted the importance of developing a support network outside the treatment setting, in order that the offender can explain his risk factors and relapse prevention plans to these people so they can offer appropriate support. In addition, a support network would also provide an opportunity for the client to enhance his social skills. Environmental factors that influenced the treatment outcome were the lack of facilities and organisations available in the community, for adults with intellectual disabilities. This presented an obstacle in

finding the client with a means of a suitable social outlet. Whilst a number of organisations were approached only one was willing to assess Mr J. for eligibility. This was for a scheme organised by Birmingham City Council, whereby individuals with intellectual and physical disabilities are paired up with a volunteer to befriend and provide social support to them. However, this scheme required the client being registered with a social worker, and so an initial referral to social services by his probation officer was needed before the eligibility assessment could be made. This was waiting to be completed whilst the intervention sessions were being conducted.

The treatment format in this case study lacked the advantages of a group intervention setting, whereby other group members can provide a means of observing and practising skills, as well as to confront or offer additional support to each other. Group interventions have also generally been argued to be more efficacious than individual work with sex offenders, since secrecy is integral to the offender's methodology in committing offences and so potentially easier to manipulate an individual therapist (Rose et al., 2002). Unfortunately, it should be noted that Mr J. had previously not responded well to group intervention or made the kind of gains that other members of this adapted group had been able to achieve, and which was why he was subsequently referred for individual work. As such, this work was intended to further enhance his social skills, ameliorating some identified deficits with regard to this dynamic risk domain.

3.2 Client characteristics and assessment issues

A significant attribute of the treatment outcome of the case study was the client's total lack of social support. The client continued to spend the majority of his time on his own, limiting the possibility of practising his newly learnt skills and establishing friendships.

The client was compliant and engaged well during the assessment intervention phase. However, a major limitation of the assessment was the reliance on the client's self-report as a measure of his success with his social skills within the community. The client also presented an excessive desire to please the therapist, which at times was detrimental to achieving open, objective and honest responses and hindering the reliability of the data. It was a challenge to reduce the client's tendency to suggestible responding. The use of continual positive appraisal was provided throughout the

intervention and did appear to reduce this tendency somewhat, as he began to feel more confident.

An important resilience factor was the client's willingness, enthusiasm and motivation for the intervention, although this may have been a reflection of his tendency to want to please. Dependency issues were however apparent during the course of the intervention, and presented a real concern that the client would rely upon the sessions as a means of social interaction to relieve his loneliness and act as a substitute friendship. Clear boundaries had to be expressed in a sensitive manner throughout. Mr J. did eventually come to understand the difference between a professional and a potential friend.

Cognitive and memory deficits also had a negative effect on the treatment outcome. Despite a simple format using pictorial aids and the use of repetition, Mr J. had difficulty recalling past experiences and the work done, even at the end of the sessions. This limited the information gained and the accuracy of events reported by the client. Furthermore, there was no third party information i.e., family members to verify or add additional information to the client, as he had no family contact. Obtaining relevant historical and background information was therefore problematic, resulting in gaps in the knowledge and understanding of the client's problems, making formulation difficult.

In relation to the assessment process of individuals with intellectual difficulties, a number of issues were apparent. Deficits in language skills and cognitive processes can lead to problems with the comprehension of complex language (Boer et al.1995), and this can affect the validity and reliability of the assessment. It was recognised that memory difficulties, acquiescence and suggestible responding were evident during the administration of tests. Such responding has been previously cited in the literature as an inherent problem in the assessment of intellectually disabled offenders, and this is known to impact negatively upon the validity of self-report measures (Clare, 1993). Most tests are deemed unsuitable for individuals with mild intellectual disabilities, due to the comprehension difficulties and literacy deficits and the lack of normative data provided for this population (Clare, 1993). There are also very few tests which can be used with offenders with lower levels of intellectual functioning (Keeling, Rose & Beech, 2007) and so this case study made use of only a small number of psychometric tests.

Psychometric tests tend to be standardised on mainstream populations, limiting the generalisibility of findings. The full battery of adapted STEP tests used within the Probation Service is currently awaiting normative data and so scores could not be accurately interpreted. Recently, since the completion of the treatment in this case study, some of the adapted psychometric battery such as the Adapted Victim Empathy Consequences Task and the Adapted Relapse Prevention Interview have now been demonstrated to have reasonable psychometric properties as determined by internal consistency and factor analysis, (Williams, Wakeling & Webster, 2007). A failure to provide test-retest reliability data also prevented clinical change from being calculated, only a visual inspection of the items was possible. Adapting current reliable psychological tests for individuals with intellectual disabilities is an acceptable method (Clare, 1993).

Only recently have researchers examined the usefulness of psychometric tests for this population, for instance Keeling, Beech and Rose (2007) recommended an adapted version of the Social Intimacy Scale (Miller & Lefcourt, 1982) and the Victim Empathy Distortion Scale (Beckett & Fisher, 1994) for use with intellectually disabled sex offenders, as both tests have revealed good internal consistency and test-retest reliability. Unfortunately these tests were not available at the time this case study was in progress. Keeling, Beech and Rose (2007) also investigated the psychometric properties of other adapted psychometric tests, such as the Criminal Sentiments Scale (Gendreau et al., 1979) and the Relationship Scale Questionnaire (Griffin & Bartholomew, 1994), and found that these tests had limited utility for this population.

Despite attempts to simplify the content and language to improve readability of the existing psychometric tests, when applying psychometrics which have no intellectually disabled normative data available a degree of caution is required until the psychometric qualities have been shown to be successful. It should be noted that the validity of the data and an accurate interpretation of the psychometric results within this case study may have been hindered, given that the psychometric tests used were not standardised on an intellectually disabled population.

3.3 Treatment length

The duration of the intervention was very short and may have negatively affected the treatment outcome. It is likely that the client would have benefited from a longer

intervention, in order to help the client consolidate the information and provide more time in which to practise the skills learnt.

Research suggests that longer periods of treatment are needed for the intellectually disabled group. In a comparison of the treatment length of intellectually disabled men who were sentenced to one or two years probation for sex offences, it was found that sex offenders that received less than one year of treatment had significantly poorer progress and more pro-offending cognitive distortions and had a greater likelihood of reoffending compared to those treated for at least two years (Lindsay & Smith, 1998). In contrast, Rose et al. (2002) reported treatment success after only a 16 week intervention. The duration of the programme may therefore only be partially accountable for Mr J.'s poor performance on the ASOTP and the rate of his success on the current intervention.

3.4 Risk assessment

The issues associated with the assessment of risk in this case study need to be highlighted. This case study used both static and dynamic actuarial measures that had not been specifically normed on the population with intellectual disabilities. Research from the previous chapter suggested the Assessment of Risk Manageability for Intellectually Disabled Individuals who Offend (ARMIDILO; Boer, Tough & Haaven, 2004), was better able to distinguish between recidivist and non-recidivists and had a large effect size.

The ARMIDILO was particularly useful as it helped inform treatment goals. The intervention strategies were designed specially to reduce the client's acute dynamic risk factors by improving his relationship skills and general coping strategies, and aimed to make positive changes in his social relationships and time management skills by promoting the use of leisure time.

Research from the previous chapter also indicated that static risk instruments had poor predictive validity for offenders with special needs. For instance, the RRASOR produced an AUC of .53 for special needs offenders, which was little better than chance level. In terms of practical utility the RRASOR did not provide a risk label, only a score with the assumption of additivity (i.e., the higher the score equals higher risk), which was not particularly helpful for a clinician assessing the level of risk posed. However, the RRASOR was initially intended to be used only to screen offenders into relative risk levels and these risk levels then adjusted by the

consideration of other relevant information such as deviant sexual preferences and treatment compliance. A further problem encountered with static actuarial instruments was that the scoring instructions set out in RM2000, did not allow for self-reported offences, only criminal appearances in court and therefore under-estimated the level of risk. Strict adherence to the scoring procedures would not be an accurate reflection of risk, particularly in light of the knowledge that legal prosecutions are less common for sexually offending men who have intellectual disabilities (Swanson & Garwick, 1990). The RM2000 was however selected as it is the risk assessment used within West Midlands Probation Services and it has had some empirical support, approaching moderate predictive accuracy for sexual recidivism with an AUC of 0.63 (Wilcox, Beech, Markall & Blacker, 2009). More empirical research is required before a consensus can be made regarding the applicability of mainstream actuarial instruments to special needs sexual offenders.

Final note

Throughout the course of the intervention the client presented as highly motivated, enthusiastic and willing to engage. A function of the treatment outcome could be attributable to multiple factors such as poor self-efficacy, a lack of social support and environmental factors. An important consideration during the intervention was the issue of therapist dependency.

This case study illustrated some of the difficulties in assessment and interventions for learning disabled clients. A limitation of this case study was the availability of adequate psychometric assessments that have been appropriately normed on an intellectually disabled sample.

Chapter Six

Discussion

This thesis has examined risk assessment instruments used in the prediction of risk in adult sexual offenders. A number of different risk assessment instruments have been developed by researchers, some of which are more widely available to clinicians. Evaluation of the effectiveness of these tools by researchers has shown support for the majority of risk assessment instruments, in their ability to accurately distinguish between recidivists and non-recidivists.

Within the systematic literature review it was identified that such research has predominantly been based on mainstream sex offender populations, with limited application to subgroups of sexual offenders such as those with intellectual disabilities, whose characteristics and offending patterns can be markedly different from mainstream offenders. A lower propensity towards criminal convictions, a tendency towards male victims and an inability to maintain stable and lasting relationships have been shown to be features of the intellectually disabled sexual offender population, all of which would lead to an inaccurate assessment of risk based on static actuarial measures.

When examining the research studies on the predictive validity of risk assessment instruments based on mainstream samples, a number of methodological limitations in terms of measurement and selection bias were recognised within the systematic review of the literature. Variations in the duration of the follow-up periods used within the studies, as well as the definitions of sexual recidivism and the types of sexual offences included within the definitions, made comparisons between different studies' findings problematic as they were not all measuring the outcome. Similarly, different types of methods to calculate inter-rater reliability hindered meaningful comparisons, and the general failure of most studies to include inter-rater reliability data was a criticism within this field of research. An overall lack of clarity within the reporting and a failure to state the level of the assessor training has impacted on the quality of previous research in terms of the reliability and validity of findings.

In an attempt to address the lack of research on the predictive validity of risk assessment instruments with an intellectually disabled sexual offender population, four risk assessment instruments, the RRASOR, SVR-20, RM2000-V and the

ARMIDILO Stable, and Acute dynamic subscales were applied and cross validated to 44 special needs offenders and compared to a matched sample of 44 mainstream offenders. The participants in the two groups were matched on all the risk items within the RRASOR tool and so posed the same actuarial level of risk, yet the special needs group the RRASOR performed little better than chance (AUC = .53) with similar results for the RM2000-V (AUC = .50).

Using a lengthy mean follow-up period of 8.8 years the results of this study found that the ARMIDILO instrument was the best predictor for sexual reconviction among offenders with special needs (ARMIDILO-Stable, AUC = .60; ARMIDILO-Acute, AUC = .73). Using unofficial data sources for sexual recidivism behaviour, the ARMIDILO-Acute scale achieved an AUC of .76, which is above the upper limits of accuracy recommended by Janus and Meehl (1997) for the sexual offender commitment selection process. The predictions using the SVR-20 yielded a greater score (AUC =.73) for the mainstream sample, in comparison to the intellectually disabled sample (AUC =.45).

Offenders within the special needs group all attended an adapted sex offender treatment programme, based on the needs of lower functioning men and comprised of those who meet the classification for an intellectual disability and those who were borderline functioning. Preliminary analysis within the special needs group, indicated that the ARMIDILO-Acute, SVR-20 Psychosocial Affect, and Overall scales were better predictors of sexual recidivism for the intellectually disabled subgroup (AUCs ranging from .75 to .88), compared to the borderline subgroup, where these tools performed at chance level.

The findings are consistent with previous research by Wilcox et al. (2009) in that both studies suggest that static risk instruments were of limited value for this special needs population. This study also highlights the value of dynamic risk factors in the assessment of sexual risk.

This research represented a fairly large sample of special needs offenders and attempted to include all offenders who participated in the community sex offender treatment programme, thus limiting the selection bias observed in other research on risk assessment instruments. By using official reconviction data in conjunction with sexual recidivism data based on unofficial data sources and a broad definition of sexual recidivism, it was hoped that this would overcome issues relating to using only

sexual reconviction data, and the problems associated with obtaining a prosecution when an offender has an intellectual disability.

Previous research has focused exclusively on 'risk' items, and in an attempt to overcome this criticism, this study encompassed protective factors that reduce the level of risk. The scoring system for the ARMIDILO incorporated protective factors into the risk decision, for example, an offender may have learnt positive coping strategies which may be judged as a protective factor reducing rather than increasing his risk. Rogers (2000) suggests that this type of rating produces a more balanced assessment, as there is not an over focus on risk, thus reducing professional negativism and client stigmatisation.

The RRASOR, a commonly used static risk instrument was examined in detail. The RRASOR appeared to be remarkably robust, with good empirical foundation and a large test construction sample. Static tools such as the RRASOR can easily be scored from administrative records making it time efficient, in comparison to dynamic risk tools or instruments that work along the lines of structured clinical judgment such as the SVR-20, which require more information to complete and information from a variety of different sources. A number of cross-validation studies found in the literature reviewed on the RRASOR, suggested that it held moderate and consistent levels of predictive validity across mainstream offender samples and has yielded good inter-rater reliability across a number of studies. The accuracy of the RRASOR has been shown to vary across ethnicity and offender subgroups, and like many other tools currently available lacked research using populations of offenders with intellectual disabilities.

A limitation of the RRASOR is its exclusive inclusion of only static risk factors in the assessment of risk. Static risk variables do not allow change to be measured, or reflect fluctuations in risk and so an offender may make many positive gains after receiving treatment but would remain at the same level of risk. This instrument also holds similar criticisms identified by Rogers (2000), in that it is subject to the problems with multi-collinearity, as it is based on the assumption of additivity, for instance that more indicators equates to more risks. The RRASOR does not hold good grounds in legal proceedings and therefore it is necessary that both researchers and clinicians continue to scrutinize the application of this tool.

In terms of the context of the 'What Works' approach (McGuire & Priestley, 1995) and the 'risk, needs and responsivity' principles (Andrews & Bonta, 2003), the

research has attempted to examine and addressed issued relating to risk for intellectually disabled sexual offenders and the case study has highlighted some of the needs issues relating the assessment approach for this population. The research conducted within this thesis suggests that the predictive validity of static actuarial tools are not as effective in predicting the level of risk, and that a dynamic risk assessment tool was of more potential value for this population of sexual offenders. The dynamic risk instrument, the ARMIDILO was also used within the case study.

The case study illustrates risk assessment in practice and how risk assessment tools can aid in the treatment formulation, by targeting dynamic risk factors. In this case the ARMIDILO tool helped inform treatment needs. Given that static risk instruments provide a probabilistic risk level that is not amendable to change, dynamic risk instruments are particularly important in providing guidance on the treatment and management of sexual offenders. Both the SVR-20 and the ARMIDILO tool have the dynamic risk variable 'relationship problems'. A failure to establish or maintain stable relationships has been found to be associated with sexual violence and may restrict a person's opportunity for appropriate sexual relations (Hanson, 1997; Hanson & Bussière, 1996). Similarly, social skills deficits have been associated with sexual offending behaviour (Marshall, 2001). Research has noted that individuals with intellectual disabilities have poor social functioning (Griffiths, Hindsburger & Christian, 1985) and a greater tendency towards social skills deficits and suffer high levels of loneliness, and low self-esteem (Boer et al., 1995).

This case study illustrates a treatment intervention to address a social skills deficit. A formulation based on risk assessment, interviews and psychometric testing all suggested that the client, an intellectually disabled sexual offender, would benefit from a social skills intervention to increase his social skills, general coping and self-esteem. Previously, facilitators on an Adapted Sex Offender Treatment Programme had expressed concerns that this client had not benefited from treatment and that his level of risk remained that same.

This individualised social skills programme was designed to focus specifically on dynamic risk factors, however, the pre-post treatment outcome measures showed no real significant treatment gains or clinical change. Some small positive changes were noticed in regards his social functioning ability such as his level of eye contact, however in terms of the level of dynamic risk this did not shift significantly within the short time frame of the intervention. Despite the client being highly motivated to

reduce his emotional loneliness and social isolation, he became dependant on the therapist as a substitute friend. Issues of therapist dependency were an impediment that continually had to be addressed throughout the treatment. Other client characteristics that influenced the treatment outcome was the client's difficulties in retaining information presented to him during the sessions.

Limitations with this thesis

The first chapter examined the predictive validity of actuarial risk assessment tools using a systematic approach. Time constraints, a language barrier and resources led to the exclusion of unpublished studies, resulting in a publication bias. Ideally a more comprehensive review with a wider sample of databases might have obtained missing unpublished research and included different languages in the quality assessment.

A major limitation within this research study was that formal assessments of IQ were not conducted in a standardised fashion at the time the special needs treatment group was being developed and held, as such not all offenders had IQ tests, therefore the analysis of those offenders who had been formally tested was based on a very small sample. The findings within this research therefore need to been viewed with some discretion, and replication of the results is necessary before firm conclusions can be drawn.

Another shortcoming of this research was that whilst a number of risk assessment instruments were completed based upon strict adherence to the manual's instructions, and upon the guidance from clinical psychologist on how to complete the risk assessment tools, no formal training was provided. A lack of inter-rater reliability in previous studies examining the predictive validity of risk assessment tools was a criticism that was highlighted within the systematic review. However, it was not possible in the research chapter to obtain inter-rater reliability as another rater was not available.

Recommendations for future research

Replication of this research on a larger sample base, using a prospective design to avoid the likelihood of predictor-criterion contamination is required. Similarly, the incorporation of multiple sources of information such as psychometric testing and the inclusion of interview-based sources of information in future studies would increase the reliability of the data and would be of added value, given that previous research

has over-relied on archival file coding. Future research would also be advised to assess intelligence across all participants in a standardised and formal manner.

Risk assessment tools such as the RRASOR have clear benefits for busy clinicians as they are time efficient to administer, however, the RRASOR had no available manual for instructing its correct usage. It some instances it may be prudent and necessary to use a modified scoring system, in order to allow for the inclusion of self-reported offences or those offences that have been known to have been committed but did not amount to a prosecution (a common theme in sexual offenders with intellectual disabilities). When a modified version of a risk assessment tool is used however, the results cannot easily be generalised unless all studies use the same scoring procedures.

Although it was not possible within the scope of this study, it would be advantageous to examine the mediating variables such as isolation or medication, and the moderator variables such as ethnicity or treatment length when examining the predictive validity of risk assessment instruments.

Future researchers who are interested in examining offenders with intellectual disabilities may select and obtain their sample on the basis of all offenders who have participated in an adapted treatment programme. Such treatment programmes may contain offenders with a wide range of intellectual functioning and so it would be advisable for future researchers to make a distinction between offenders with intellectual disabilities (i.e., those with an IQ below 75) and those who fall within the borderline range when examining offenders, as the failure to differentiate between the two may confound and skew the results. The findings within this thesis were valuable, however, being based on a very small sample of borderline offenders the results can only be viewed with caution and seen as preliminary. More research is required to examine risk assessment tools with borderline IQ offenders.

Conclusions

This thesis has examined an under researched and valuable topic and had furthered our knowledge on the use of risk assessment tools with intellectually disabled populations. It is paramount that clinicians are aware of the limits of risk instruments when assessing the level of risk. The findings of this thesis suggest that risk assessment instruments developed on mainstream offender populations are not as

accurate in their predictions when applied to lower intellectually functioning offenders.

This research has shown the potential for the clinical application of the ARMIDILO tool which was specifically designed for individuals with intellectual disabilities, although further validation is necessary to support its use in practice. It is also suggested that practitioners target their treatments to the idiosyncratic dynamic risk factors of offenders in order to reduce future recidivism. It is hoped that the findings are of value to clinicians and researchers in the future.

References

*References excluded from systematic review
**Unallocated studies in the systematic review.

Allan, A., Dawson, D., & Allan, M. M. (2006). Prediction of the risk of male sexual reoffending in Australia. *Australian Psychologist, 41*, 60-68.

Allan, M., Grace, R. C., Rutherford, B., & Hudson, S. M. (2007). Psychometric assessment of dynamic risk factors for child molesters. *Sexual Abuse: A Journal of Research and Treatment, 19*, 347-367.

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders – Text revision (DSM-TR).* Washington, DC: American Psychiatric Association.

American Psychological Association. (1992). Ethical principles of psychologists and code of conduct. *American Psychologist, 47*, 1597-1611.

Ashman, L., & Duggan, L. (2003). *Interventions for learning disabled sex offenders (Cochrane Review).* The Cochrane Library 2003, Issue 1. Oxford, UK: Update Software.

*Austin, J., & Johnson, K. D. (2003). *Reliability and validity study of the Static-99/RRASOR sex offender risk assessment instruments.* The Institute on Crime, Justice and Corrections, George Washington University.

Bagley, C., & Prichard, C. (2000). Criminality and violence in intra- and extra-familial child abusers in a 2-year cohort of convicted perpetrators. *Child Abuse Review, 9*, 264-274.

*Barbaree, H. E., Langton, C. M., & Peacock, E. J. (2006a). Different actuarial risk measures produce different risk rankings for sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 18*, 423-440.

*Barbaree, H. E., Langton, C. M., & Peacock, E. J. (2006b). The factor structure of static actuarial items: Its relation to prediction. *Sexual Abuse: Journal of Research and Treatment, 18*, 207-226.

Barbaree, H.E., Seto, M.C., Langton, C.M., & Peacock, E.J. (2001). Evaluating the predictive accuracy of six risk assessments for adult sex offenders. *Criminal Justice and Behavior, 28,* 490-521.

Bartosh, D.L., Garby, T., Lewis, D. & Gray, S. (2003). Differences in the predictive validity of actuarial risk assessments in relation to sex offender types. *International Journal of Offender Therapy and Comparative Criminality, 47,* 422-438.

*Bates, A., Falshaw, L., Corbett, C., Patel, V., & Friendship, C. (2004). A follow-up study of sex offenders treated by Thames Valley sex offender groupwork programme, 1995-1999. *Journal of Sexual Aggression, 10*, 29-38.

Beech, A.R., Fisher, D.D., & Beckett, R. (1998). *Step 3: An evaluation of the Prison Sex Offender Treatment Programme*. A report for the Home Office by the STEP team. London: Home Office.

Beech, A.R., Fisher, D.D., & Thornton, D. (2003). Risk assessment of sex offenders. *Professional Psychology: Research and Practice*, *34,* 339-352.

*Beech, A., & Ford, H. (2006). The relationship between risk, deviance, treatment outcome and sexual reconviction in a sample of child sexual abusers completing residential treatment for their offending. *Psychology, Crime and Law, 12*, 685-701.

Beech, A., Friendship, C., Erikson, M., & Hanson, R. K. (2002). Relationship between static and dynamic risk factors and reconviction in a sample of U.K. child abusers. *Sexual Abuse: A Journal of Research and Treatment, 14*, 155-167.

Beech, A., & Ward, T. (2004). The integration of etiology and risk in sex offenders: A theoretical framework. *Aggression and Violent Behavior, 10,* 31-63.

Bellack, A., & Hersen, M. (1998). *Behavioural assessment. A practical handbook.* London: Allyn and Bacon.

Bengtson, S., & Långström, N. (2007). Unguided clinical and actuarial assessment of re-offending risk: A direct comparison with sex offenders in Denmark. *Sexual Abuse: A Journal of Research and Treatment, 19,* 135-153.

Bengtson, S. (2008). Is newer better? A cross-validation of the static-2002 and the Risk Matrix 2000 in a Dutch sample of sexual offenders. *Psychology, Crime and Law, 14,* 85-106.

*Berlin, F. S., Galbreath, N. W., Geary, B., & McGlone, G. (2003). The use of actuarials at civil commitment hearings to predict the likelihood of future sexual violence. *Sexual Abuse: Journal of Research and Treatment, 15,* 377-382.

Boer, D.P. (2006). Sex offender risk assessment strategies: is there a convergence of opinion yet? *Sexual Offender Treatment, 1,* 1-4.

Boer, D.P., Gauthier, C.M., Watson, D.R., Dorward, J., & Kolton, D.J. (1995). The assessment and treatment of intellectually disabled sex offenders: The Regional Psychiatric Centre (Pacific) "Northstar" treatment and relapse prevention program. *Paper presented at Towards a National Strategy: A Conference on Intervention with Sex offenders, Toronto, Canada.*

Boer, D.P., Hart, S.D., Kropp, P.R., & Webster, C.D. (1997). *Manual for the Sexual Violence Risk-20 (SVR-20): Professional guidelines for assessing risk of sexual violence.* Vancouver, Canada: British Columbia Institute Against Family Violence.

*Boer, D. P., Tough, S., & Haaven, J. (2004). Assessment of risk manageability of intellectually disabled sex offenders. *Journal of Applied Research in Intellectual Disabilities, 17,* 275-283.

Bohrnstedt, G. W., & Knoke, D. (1994). *Statistics for social data analysis* (3rd ed). Itasca, IL: F. E. Peacock Publishers.

Borum, R. (1996). Improving the clinical practice of violence risk assessment: Technology, guidelines and training. *American Psychologist, 51*, 945-956.

Bregman, S. (1984). Assertiveness training for mentally retarded adults. *Mental Retardation, 22*, 12-16.

British Psychological Society (2001). *Learning Disability: Definitions and Contexts.* Leicester: Professional Affairs Board of the British Psychological Society.

Brown, H., & Stein, J. (1997). Sexual abuse perpetrated by men with intellectual disabilities: A comparative study. *Journal of Intellectual Disability Research, 41*, 215-224.

Brown, H., & Thompson, D. (1997). The ethics of research with men who have learning disabilities and abusive sexual behaviour: a minefield in a vacuum. *Disability and Society, 12*, 695-707.

**Brown, J. A. (2004). A comparison of actuarial methods of predicting sexual dangerousness. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 65*(2-B), 1068.

Burgess, R., Jewitt, R., Sandham, J., & Hudson, B. (1980). Work with sex offenders: A social skills training group. *British Journal of Social Work, 10*, 133-142.

Campbell, T.W. (2000). Sexual predator evaluations and phrenology: Considering issues of evidentiary reliability. *Behavioral Sciences and the Law, 18,* 111-130.

*Campbell, T.W. (2003). Sex offenders and actuarial risk assessments: Ethical considerations. *Behavioral Sciences and the Law, 21*, 269-279.

**Caperton, J. D. (2006). Predicting recidivism among sex offenders: Utility of the static-99, minnesota sex offender screening tool-revised, and psychopathy checklist-revised. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 66*(9-B), 5080.

Christensen, I. & Mendoza, J.I. (1986). A method of addressing change in a single subject: An alteration of the RC index. *Behavior Therapy, 17*, 305-308.

Clare, I.C.H. (1993). Issues in the assessment and treatment of male sex offenders with mild learning disabilities. *Sexual and Marital Therapy, 8*, 167-80.

Coleman, E.M., & Haaven, J. (2001). Assessment and treatment of intellectually disabled sexual abusers. In M.S. Carich & S.E. Mussack (Eds.), *Handbook of sexual abuser assessment and treatment* (pp. 193-210). Brandon, VT: Safer Society Press.

Courtney, J., & Rose, J. (2004). The effectiveness of treatment for male sex offenders with learning disabilities: A review of the literature. *Journal of Sexual Aggression, 10,* 215-236.

Craig, L. A., Beech, A., & Browne, K. D. (2006a). Sexual recidivism: Advances in actuarial estimates. *Forensic Updates, 85,* 13-18.

Craig, L. A., Beech, A., & Browne, K. D. (2006b). Cross-validation of the risk matrix 2000 sexual and violent scales. *Journal of Interpersonal Violence, 21,* 612-633.

Craig, L. A., Browne, K. D., & Beech, A. (2008). *Assessing risk in sex offenders.* London: Wiley.

Craig, L. A., Browne, K. D., Beech, A., & Stringer, I. (2004). Personality characteristics associated with reconviction in sexual and violent offenders. *Journal of Forensic Psychiatry and Psychology, 15*, 532-551.

Craig, L. A., Browne, K. D., Beech, A., & Stringer, I. (2006a). Differences in personality and risk characteristics in sex, violent and general offenders. *Criminal Behaviour and Mental Health, 16*, 183-194.

Craig, L. A., Browne, K. D., Beech, A., & Stringer, I. (2006b). Psychosocial characteristics of sexual offenders and the relationship to sexual reconviction. *Psychology, Crime and Law, 12,* 231-243.

*Craig, L.A., Browne, K.D., & Stringer, I. (2003). Risk scales and factors predictive of sexual offence recidivism. *Trauma, Violence, and Abuse, 4*, 45-69.

*Craig, L.A., Browne, K.D., & Stringer, I. (2004). Comparing sex offender risk assessment measures on a UK sample. *International Journal of Offender Therapy and Comparative Criminology, 48*, 7-27.

Craig, L.A., Browne, K.D., Stringer, I. & Beech, A.R. (2004). Limitations in actuarial risk assessment of sexual offenders: a methodological note. *The British Journal of Forensic Practice, 1*, 16-32.

Craig, L.A., Browne, K.D., Stringer, I., & Beech, A.R. (2005). Sexual recidivism: a review of static, dynamic and actuarial predictors. *Journal of Sexual Aggression, 11*, 65-84.

*Craig, L. A., Browne, K. D., Stringer, I., & Hoguee, T. E. (2008). Sexual reconviction rates in the United Kingdom and actuarial risk estimates. *Child Abuse and Neglect, 32*, 121-138.

Craig, L., & Hutchinson, R.B. (2005). Sexual offenders with Learning disabilities: risk, recidivism and treatment. *Journal of Sexual Aggression, 11*, 289-304.

Craig, L. A., Thornton, D., Beech, A., & Browne, K. D. (2007). The relationship of statistical and psychological risk markers to sexual reconviction in child molesters. *Criminal Justice and Behavior, 34*, 314-329.

Craig, L.A., Stringer, I., & Moss, T. (2006). Treating sexual offenders with learning disabilities in the community. *International Journal of Offender Therapy and Comparative Criminology, 50*, 369-390.

Craissati, J., & Beech, A. (2005). Risk prediction and failure in a complete urban sample of sex offenders. *Journal of Forensic Psychiatry and Psychology, 16*, 24-40.

*Craissati, J., & Beech, A. (2006). The role of key developmental variables in identifying sex offenders likely to fail in the community: An enhanced risk prediction model. *Child Abuse and Neglect, 30,* 327-338.

*Craissati, J., Falla, S., McClurg, G., & Beech, A. (2002). Risk, reconviction rates and pro-offending attitudes for child molesters in a complete geographical area of London. *Journal of Sexual Aggression, 8,* 22-38.

Crawford, D.A (1981). Treatment approaches with paedophiles. In M. Cook & K. Howells (Eds.), *Adult sexual interest in children* (pp. 181-217). London: Academic Press.

Day, K. (1994). Male mentally handicapped sex offenders. *British Journal of Psychiatry, 165,* 630-639.

*Dempster, R. J., & Hart, S. D. (2002). Relative utility of fixed and variable risk factors in discriminating sexual recidivists and nonrecidivists. *Sexual Abuse: A Journal of Research and Treatment, 14,* 121-138.

de Vogel, V., de Ruiter, C., van Beek, D., & Mead, G. (2004). Predictive validity of the SVR-20 and static-99 in a Dutch sample of treated sex offenders. *Law and Human Behavior, 28,* 235-251.

Doren, D.M. (2002). *Evaluating sex offenders: A manual for civil commitment and beyond*. London: Sage.

*Doren, D.M. (2004). Stability of the interpretative risk percentages for the RRASOR and Static-99. *Sexual Abuse: A Journal of Research and Treatment, 16,* 25-36.

*Doren, D. M. (2006). What do we know about the effect of aging on recidivism risk for sexual offenders? *Sexual Abuse: A Journal of Research and Treatment, 18,* 137-157.

Dietiker, J., Dittmann, V., & Graf, M. (2007). Risk assessment of sex offenders in a german-speaking sample. applicability of PCL-SV, HCR-20+3, and SVR-20. *Nervenarzt, 78,* 53-61.

Ducro, C., & Pham, T. (2006). Evaluation of the SORAG and the static-99 on Belgian sex offenders committed to a forensic facility. *Sexual Abuse: A Journal of Research and Treatment, 18,* 15-26.

**Eher, R., Rettenberger, M., Schilling, F., & Pfafflin, F. (2008). Validity or benefit? the prediction of relapses with static-99 and SORAG. A prospective study with 275 sex offenders. *Recht and Psychiatrie, 26*(2), 79-88.

*Epperson, D. L., Kaul, J. D., Huot, S. J., Hesselton, D., Alexander, W., & Goldman, R. (1998). *Minnesota Sex Offender Screening Tool–Revised (MnSOST-R).* St. Paul: Minnesota Department of Corrections.

English, K., Retzlaff, P., & Kleinsasser, D. (2002). The Colorado sex offender risk scale. *Journal of Child Sexual Abuse, 11,* 77-96.

Falshaw, L., Friendship, C., & Bates, A. (2003). *Sexual offenders-measuring reconviction, reoffending and recidivism.* Research Development and Statistics Directorate: Findings Report 183. Home Office: available at http://www.homeoffice.gov.uk/rds/pdfs2/r183.pdf

*Ferguson, G. E. (1998). An investigation into the validity of the registrant risk assessment scale as a legal tool and clinical instrument. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 58*(9-B), 5184.

*Ferguson, G. E., Eidelson, R. J., & Witt, P. H. (1998). New Jersey's sex offender risk assessment scale: Preliminary validity data. *Journal of Psychiatry and Law, 26,* 327-351.

Firestone, P., Nunes, K. L., Moulden, H., Broom, I., & Bradford, J. M. (2005). Hostility and recidivism in sexual offenders. *Archives of Sexual Behavior, 34,* 277-283.

Fisher, D., & Thornton, D. (1993). Assessing risk of re-offending in sexual offenders. *Journal of Mental Health, 2,* 105-117.

Fliess, J.L. (1981). *Statistical methods for rates and proportions.* New York: Wiley.

Friendship, C., Thornton, D., Erikson, M., & Beech, A. (2001). Reconviction: A critique and comparison of two main data sources in England and Wales. *Legal and Criminological Psychology, 6,* 121-129.

Friendship, C., Beech, A.R., & Browne, K.D. (2002). Reconviction as an outcome measure in research: a methodological note. *British Journal of Criminology, 42*, 442-444.

Frisch, M.B., & Jessop, N.S. (1989). Improving WAIS-R estimates with the Shipley-Hartford and Wonderlic Personnel Tests: Need to control for reading ability. *Psychological Reports, 65*, 923-928.

*Furr, K. D. (. (1993). Prediction of sexual or violent recidivism among sexual offenders: A comparison of prediction instruments. *Sexual Abuse: A Journal of Research and Treatment, 6*, 271-286.

*Gentry, A. L., Dulmus, C. N., & Theriot, M. T. (2005). Comparing sex offender risk classification using the static-99 and LSI-R assessment instruments. *Research on Social Work Practice, 15*, 557-563.

**Gore, K. S. (2008). Adjusted actuarial assessment of sex offenders: The impact of clinical overrides on predictive accuracy. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 68*(7-B), 4824.

Gray, N.S., Fitzgerald, S., Taylor, J., MacCulloch, M.J., & Snowden, R.J. (2007). Predicting future reconviction in offenders with intellectual disabilities: the predictive efficacy of VRAG, PCL-SV, and the HCR-20. *Psychological Assessment, 19*, 474-479.

Green, G., Gray, N.S., & Willner, P. (2002). Factors associated with criminal convictions for sexually inappropriate behaviour in men with learning disabilities. *Journal of Forensic Psychiatry, 13,* 578-607.

*Green, G., Gray, N. S., & Willner, P. (2003). Management of sexually inappropriate behaviours in men with learning disabilities. *Journal of Forensic Psychiatry and Psychology, 14*, 85-110.

Gresham, F.M. (1998). Social skills Training: Should we raze, remodel or rebuild?. *Behavioral Disorders, 24*, 19-25.

Griffiths, D., Hinsburger, D., & Christian, R. (1985). Treating developmentally handicapped sexual offenders. The York Behavioural Management Services Treatment Program. *Psychiatric Aspects of Mental Retardation Reviews, 4,* 49-54.

Grove, W.M., Zald, D.H., Lebow, B.S., Snitz, B.E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment, 12,* 19-30.

Grubin, D. (1999). Actuarial and clinical assessment of risk in sex offenders. *Journal of Interpersonal Violence, 14*, 331-43.

**Haag, A. M. (2006). Do psychological interventions impact on actuarial measures: An analysis of the predictive validity of the static-99 and static-2002 on a re-conviction measure of sexual recidivism. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 66*(8-B), 4531.

Haaven, J. (2005). Treating intellectually disabled sex offenders allows us to see with 'new eyes'. *NOTANews*, 49, 2-5.

Haaven, J.L., & Coleman, E.M. (2000). Treatment of the developmentally disabled sex offender. In D.R. Laws, S.M. Hudson, & T. Ward (Eds.), *Remaking relapse prevention with sex offenders: A sourcebook* (pp. 167-186). Thousand Oaks, CA: Sage.

Haaven, J., Little, R., & Petre-Miller, D. (1990). Treating Intellectually Disabled Sex Offenders: A Modal Residential Programme. Orwell, VT: The Safer Society Press.

Hanlon, M. J., Larson, S., & Zacher, S. (1999). The Minnesota SOST and sexual reoffending in north Dakota: A retrospective study. *International Journal of Offender Therapy and Comparative Criminology, 43*, 71-77.

Hanson, R.K. (1997). *The development of a brief actuarial risk scale for sexual offense recidivism. (User Report 97-04).* Ottawa: Department of the Solicitor General of Canada. Available at www.sgc.gc.ca/epub/corr/e199704/e199704.htm

Hanson, R.K. (1998). What do we know about sex offender risk assessment? *Psychology, Public Policy and Law, 4*, 50-72.

Hanson, R. K. (2006). Does static-99 predict recidivism among older sexual offenders? *Sexual Abuse: A Journal of Research and Treatment, 18*, 343-355.

*Hanson, R.K., Broom, I., & Stephenson, M. (2004). Evaluating community sex offender treatment programs: A 12-year follow-up of 724 offenders. *Canadian Journal of Behavioural Science, 36*, 87-96.

*Hanson, R.K., & Bussière, M.T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology*, 66, 348-362.

*Hanson, R. K., & Harris, A. (2000). *Sex offender need assessment rating (SONAR): A method for measuring change in risk levels. (*Report No. 2000-01). Ottawa, ON: Department of Solicitor General Canada.

Hanson, R. K., & Harris, A. J. R. (2001). A structured approach to evaluating change among sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 13*, 105-122.

Hanson, R.K., & Hanson-Bourgon, K. (2004). *Predictors of sexual recidivism: An updated meta-analysis.* Ottawa, Canada: Public Safety and Emergency Preparedness

Canada, Corrections Research. Available at www.psepc-sppcc.gc.ca/publications/corrections/pdf/200402)_e.pdf

*Hanson, R. K., & Morton-Bourgon, K. E. (2005). The characteristics of persistent sexual offenders: A meta-analysis of recidivism studies. *Journal of Consulting and Clinical Psychology, 73*, 1154-1163.

Hanson, R. K., & Morton-Bourgon, K. (2007). *The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis.* (Report No. 2007-01). Ottawa, ON: Public Safety and Emergency Preparedness Canada.

Hanson, R.K., Morton, K.E., & Harris, A.J.R. (2003). Sexual offender recidivism risk: What we know and what we need to know. In R. Prenky, E. Janus, M. Seto, and A.W. Burgess (Eds), *Understanding and managing sexually coercive behavior. Annals of the New York Academy of Sciences,* 989, 154-166.

Hanson, R.K., & Thornton, D. (2000). Improving risk assessments for sex offenders: A comparison of three actuarial scales. *Law and Human Behavior, 24*, 119-136.

Hanson, R. K., & Thornton, D. (2003). Notes on the development of Static-2002 (Report No. 2003-01). Ottawa, Canada: Public Safety and Emergency Preparedness Canada.

*Harris, A. J. R., & Tough, S. (2004). Should actuarial risk assessments be used with sex offenders who are intellectually disabled? *Journal of Applied Research in Intellectual Disabilities, 17*, 235-241.

Harris, G.T., & Rice, M. E. (2003). Actuarial assessment of risk among sex offenders. In R. Prenky, E. Janus, M. Seto, and A.W. Burgess (Eds), *Understanding and managing sexually coercive behavior. Annals of the New York Academy of Science,* 989, 198-210.

Harris, G.T., Rice, M.E., Quinsey, V.L., Lalumière, M.L., Boer, D., & Lang, C. (2003). A multisite comparison of actuarial risk instruments for sex offenders. *Psychological Assessment,* 15, 413-425.

Hart, S., Kropp, P.R.., Laws, D.R., Klaver, J., Logan, C., & Watt, K.A. (2003). *The risk for Sexual Violence Protocol (RSVP): Structured professional guidelines for assessing risk of sexual violence*. Vancouver, BC: The Institute Against Family Violence.

Hayes, S. (1991) Sex offenders. *Australia and New Zealand Journal of Developmental Disabilities,* 17, 221-227.

Heilbrun, K. (1992). The role of psychological testing in forensic assessment. *Law and Human Behavior, 16,* 257-272.

Henderson, M. & Hollin, C. R. (1986) 'Social skills training and delinquency', in C.R. Hollin, and P. Trower (Eds.), *Handbook of social skills training, Volume 1: Applications across the life span.* Oxford: Pergamon Press.

Holden, B. (2002). Some limitations of functional analyses: treatment of problem behaviour based on the study of alternative responses. *Behavioral Interventions, 17,* 191-209.

Hollin, C., & Trower, P. (1986). *Handbook of social skills training, Volume 1*. Oxford: Pergamon Books.

Hood, R., Shute, S., Feilzer, M., & Wilcox, A. (2002). Sex offenders emerging from long-term imprisonment- a study of their long-term reconviction rates and of parole board members' judgement of their risk. *British Journal of Criminology*, 42, 371-394.

Hopko, D.R., & Hopko, S.D. (1999). What can functional analytic psychotherapy contribute to empirically-validated treatments? *Clinical Psychology and Psychotherapy, 6,* 349-356.

Howells, K., & Day, A. (2003). Readiness for anger management: clinical and theoretical issues. *Clinical Psychology Review, 23*, 319-337.

Hudson, S.M., Wales, D.S., Barker, L., & Ward, T. (2002). Dynamic risk factors: The Kia Marama evaluation. *Sexual Abuse: A Journal of Research and Treatment*, 14, 103-119.

Jacobson, N.S. Follette, W.C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy*, 15, 336-352.

Jacobson, N.S., & Traux, P. (1991) Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59,* 12-19.

Janus, E.A., & Meehl, P.E. (1997). Assessing the legal standard for predictions of dangerousness in sex offender commitment proceedings. *Psychology, Public Policy and Law, 3,* 33-64.

*Johnson, M. E. (2003). Assessing risks and needs of sex offenders in community-based treatment programs: Differentiating among sex offenders by STATIC-99 risk factors using the BASIS-A. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 64*(6-B), 2923.

**Johansen, S. H. (2007). Accuracy of predictions of sexual offense recidivism: A comparison of actuarial and clinical methods. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 68*(3-B), 1929.

Kazdin, A.E. (2003). Clinical significance: Measuring whether interventions make a difference. In A.E. Kazdin (Ed.), *Methodological Issues and Strategies in Clinical Research, 3rd Edition* (pp. 691-710). Washington, DC: American Psychological Association.

Keeling, J.A., Beech, A.R., & Rose, J.L. (2007). The assessment of sexual offenders with an intellectual disability: the current position. *Aggression and Violent Behavior, 12,* 229-241.

Keeling, J.A., Rose, J.L., & Beech, A.R. (2006). An investigation into the effectiveness of a custody-based cognitive-behavioural treatment for special needs sexual offenders. *Journal of Forensic Psychiatry and Psychology, 17,* 372-392.

Klimecki, M., Jenkinson, J., Wilson, L. (1994). A study of recidivism among offenders with an intellectual disability. Australia and New Zealand Journal of Developmental Disabilities, 19, 209-219.

Kemshall, H. (2001). *Risk assessment and management of known sexual and violent offenders: A review of current issues*. Home Office Research Study 140, London: Home Office.

Kendall P.C. (1999). Clinical significance. *Journal of Consulting and Clinical Psychology, 67,* 285-299.

Knight, R. A. D., & Thornton, D. D. (2007). *Evaluating and improving risk assessment schemes for sexual recidivism: A long-term follow-up of convicted sexual offenders*. Retrieved July 2008 from: http://www.ncjrs.gov/pdffiles1/nij/grants/217618.pdf

Lambrick, F. (2003). Issues surrounding the risk assessment of sexual offenders with an intellectual disability. *Psychiatry, Psychology and Law,* 10, 353-358.

Lambrick, F., & Glaser, W. (2004). Sex offenders with an intellectual disability. *Sexual Abuse: A Journal of Research and Treatment*, 4, 381-392.

Långström, N. (2004). Accuracy of actuarial procedures for assessment of sexual offender recidivism risk may vary across ethnicity. *Sexual Abuse: A Journal of Research and Treatment, 16*, 107-120.

*Langevin, R. (2006). An actuarial study of recidivism risk among sex killers of adults and children: Could we have identified them before it was too late? *Journal of Forensic Psychology Practice, 6*, 29-49.

**Langton, C. M. (2003). Contrasting approaches to risk assessment with adult male sexual offenders: An evaluation of recidivism prediction schemes and the utility of supplementary clinical information for enhancing predictive accuracy. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 64*(4-B), 1907.

Langton, C. M., Barbaree, H. E., Hansan, K. T., Harkins, L., & Peacock, E. J. (2007). Reliability and validity of the static-2002 among adult sexual offenders with reference to treatment status. *Criminal Justice and Behavior, 34*, 616-640.

Langton, C. M., Barbaree, H. E., Seto, M. C., Peacock, E. J., Harkins, L., & Hansen, K. T. (2007). Actuarial assessment of risk for reoffense among adult sex offenders: Evaluating the predictive accuracy of the static-2002 and five other instruments. *Criminal Justice and Behavior, 34*, 37-59.

*Levenson, J.S., & Morin, J.W. (2006). Factors predicting selection of sexually violent predators for civil commitment. *International Journal of Offender Therapy and Comparative Criminology, 50*, 609-629.

*Levenson, J.S. (2004). Sexual predictor civil commitment: A comparison of selected and released offenders. *International Journal of Offender Therapy and Comparative Criminology, 48*, 638-648.

Lindsay, W.R. (2002). Research and literature on sex offenders with intellectual and developmental disabilities. *Journal of Intellectual Disability Research, 46*, 74-85.

Lindsay, W.R. (2004). Sex offenders: conceptualisation of the issues, services, treatment and management. In W.L. Lindsay, J.L. Taylor & P. Sturmey (Eds), *Offenders with developmental disabilities* (pp. 163-185). Chichester: Wiley

Lindsay, W.R., & Beail, N. (2004). Risk assessment: actuarial prediction and clinical judgement of offending incidents and behaviour for intellectual disability services. *Journal of Applied Research in Intellectual Disabilities*, 17, 229-234.

Lindsay, W.R., Elliot, S.F., & Astell, A. (2004). Predictors of sexual offence recidivism in offenders with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities*, 17, 299-305.

*Lindsay, W. R., Hogue, T., Taylor, J. L., Mooney, P., Steptoe, L., Johnston, S., O'Brien, G. & Smith, A.H.W. (2006). Two studies on the prevalence and validity of personality disorder in three forensic intellectual disability samples. *Journal of Forensic Psychiatry and Psychology, 17*, 485-506.

*Lindsay, W. R., Hogue, T. E., Taylor, J. L., Steptoe, L., Mooney, P., O'Brien, G., Johnston, S., & Smith, A.H.W. (2008). Risk assessment in offenders with intellectual disability: A comparison across three levels of security. *International Journal of Offender Therapy and Comparative Criminology, 52*, 90-111.

Lindsay, W.R., Law, J., Quinn, K., Smart, N., & Smith, A.H. (2001). A comparison of physical and sexual abuse: Histories of sexual and non-sexual offenders with learning disabilities. *Child Abuse and Neglect, 25,* 989-995.

Lindsay, W.R., Marshall, I., Neilson, C., Quinn, K., & Smith, H.W. (1998). The treatment of men with a learning disability convicted of exhibitionism. *Research in Developmental Disabilities*, 4, 295-316.

*Lindsay, W. R., Murphy, L., Smith, G., Murphy, D., Edwards, Z., Chittock, C., Grieve, A., & Young, S.J. (2004). The dynamic risk assessment and management system: An assessment of immediate risk of violence for individuals with offending and challenging behaviour. *Journal of Applied Research in Intellectual Disabilities, 17*, 267-274.

Lindsay, W.R., Neilson, C.Q., Morrison, F., & Smith, A.H.W. (1998). The treatment of six men with a learning disability convicted of sex offences with children. *British Journal of Clinical Psychology*, 37, 83-98.

Lindsay, W.R., Olley, S., Baillie, N., & Smith, A.H.W. (1999). The treatment of adolescent sex offenders with intellectual disability. *Mental Retardation,* 37, 320-33.

Lindsay, W.R. & Smith, A.H.W. (1998). Responses to treatment for sex offenders with intellectual disability: a comparison of men with one and two year probation sentences. *Journal of Intellectual Disability Research,* 42, 346-53.

Litwack, T. R. (2001). Actuarial versus clinical assessments of dangerousness. *Psychology, Public Policy, and Law, 7,* 409- 443.

Looman, J. (2006). Comparison of two risk assessment instruments for sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 18*, 193-206.

Lund, C. A. (2000). Predictors of sexual recidivism: Did meta-analysis clarify the role and relevance of denial? *Sexual Abuse: A Journal of Research and Treatment, 12,* 275-287.

Marshall, W.L. (2001). Enhancing social relationship skills. In M.S. Carich, & S.E. Mussack (Eds.), *Handbook for sexual abuser assessment and treatment* (pp,149-162). Brandon, VT: Safer Society Foundation.

Marshall, W.L., Barbaree, H.E., & Fernandez, Y.M. (1995). Some aspects pf social competence in sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 7*, 113-127.

Marshall, W.L., Bruce, P., Hudson, S.M., Ward, T., & Moth, B. (1996). The enhancement of intimacy and the reduction of loneliness among child molesters. *Journal of Family Violence, 11*, 219-225.

Macpherson, G. J. (1997). Psychology and risk assessment. *British Journal of Clinical Psychology, 36,* 643–645.

*Macpherson, G. J. D. (2003). Predicting escalation in sexually violent recidivism: Use of the SVR-20 and PCL: SV to predict outcome with non-contact recidivists and contact recidivists. *Journal of Forensic Psychiatry and Psychology, 14*, 615-627.

McFall, R.M. (1990). The enhancement of social skills: An information processing analysis. In W.L. Marshall, D.R Laws, & H.E. Barbaree (Eds.), *Handbook of sexual assault: Issues, theories and treatment of the offender* (pp. 311-330). New York: Plenum Press.

Mossman, D. (1994). Assessing predictions of violence: Being accurate about accuracy. *Journal of Consulting and Clinical Psychology, 62,* 783-792.

*Mossman, D. (2006). Another look at interpreting risk categories. *Sexual Abuse: A Journal of Research and Treatment, 18*, 41-63.

**Mueller, P. T. (2007). An examination of the relationship between sexual interest and risk for recidivism utilizing the abel assessment "for sexual interest(tm)" and the STATIC-99 in a sample of adult male sexual offenders. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 68*(3-B), 1937.

Mulloy, R., & Marshall, W. (1999). Social functioning. In W.L. Marshall, D. Anderson, & Y. Fernandez (Eds.). *Cognitive behavioural treatment of sexual offenders* (pp. 93-109). Chichester: Wiley.

*Nunes, K. L., & Cortoni, F. (2008). Dropout from sex-offender treatment and dimensions of risk of sexual recidivism. *Criminal Justice and Behavior, 35*, 24-33.

*Nunes, K.L., Firestone, P., & Baldwin, M.K. (2007). Indirect assessment of cognitions of child sexual abusers with implicit association test. *Criminal Justice and Behavior, 34,* 454-475.

Nunes, K. L., Firestone, P., Bradford, J. M., Greenberg, D. M., & Broom, I. (2002). A comparison of modified versions of the Static-99 and the sex offender risk appraisal guide. *Sexual Abuse: A Journal of Research and Treatment, 14*, 253-269.

*Nunes, K. L., Firestone, P., Wexler, A. F., Jensen, T. L., & Bradford, J. M. (2007). Incarceration and recidivism among sexual offenders. *Law and Human Behavior, 31*, 305-318.

*Nunes, K. L., Hanson, R. K., Firestone, P., Moulden, H. M., Greenberg, D. M., & Bradford, J. M. (2007). Denial predicts recidivism for some sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 19*, 91-105.

O'Connor, W. (1996). A problem-solving intervention for sex offenders with an intellectual disability. *Journal of Intellectual and Developmental Disability*, 21, 219-35.

O'Connor, W. (1997). Towards an environmental perspective on intervention for problem sexual behaviour in people with an intellectual disability. *Journal of Applied Research in Intellectual Disabilities*, 10, 159-175.

**Olver, M. E. (2004). The development and validation of the violence risk scale: Sexual offender version (VRS:SO) and its relationship to psychopathy and treatment attrition. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 64*(9-B), 4628.

Olver, M. E., Wong, S. C. P., Nicholaichuk, T., & Gordon, A. (2007). The validity and reliability of the violence risk scale-sexual offender version: Assessing sex offender risk and evaluating therapeutic change. *Psychological Assessment, 19*, 318-329.

Overholser, J.C., & Beck, S. (1986). Multimethod assessment of rapists, child molesters and three control groups on behavioural and psychological measures. *Journal of Consulting and Clinical Psychology, 54*, 682-687.

Paulhus, D.L. (1984). Two component model of socially desirable responding. *Journal of Personality and Social Psychology, 46,* 598-609.

Prentky, R., Harris, B., Frizell, K., & Righthand, S. (2000). An actuarial procedure for assessing risk with juvenile sex offenders. *Sexual Abuse: A Journal of Research and Treatment, 12,* 71-93.

*Proulx, J., Pellerin, B., Paradis, Y., McKibben, A., Aubut, J., & Ouimet, M. (1997). Static and dynamic predictors of recidivism in sexual aggressors. *Sexual Abuse: A Journal of Research and Treatment, 9*, 7-27.

Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (1998). *Violent offenders: Appraising and managing risk.* Washington, DC: American Psychological Association.

Rathus, S.A. (1973). A 30-item schedule for assessing assertive behaviour. *Behavior Therapy, 4,* 398-406.

Rice, M.E., & Harris, G.T. (1995). Violent recidivism: Assessing predictive validity. *Journal of Consulting and Clinical Psychology, 63,* 737-748.

*Rice, M. E., & Harris, G. T. (1997). Cross-validation and extension of the violence risk appraisal guide for child molesters and rapists. *Law and Human Behavior, 21*, 231-241.

Rice, M. E., & Harris, G. T. (2002). Men who molest their sexually immature daughters: Is a special explanation required? *Journal of Abnormal Psychology, 111*, 329-339.

Rice, M.E., & Harris, G.T. (2005). Comparing effects sizes in follow-up studies: ROC area, Cohen's *d*, and *r*. *Law and Human Behavior, 29,* 615-620.

*Roberts, C. F., Doren, D. M., & Thornton, D. (2002). Dimensions associated with assessments of sex offender recidivism risk. *Criminal Justice and Behavior, 29*, 569-589.

Robertson, I., Richardson, A.M., & Youngson, S.C. (1984). Social skills training with mentally handicapped people: a review. *British Journal of Clinical Psychology, 23*, 241-264.

Rogers, R. (2000). The uncritical acceptance of risk assessment in forensic practice. *Law and Human Behavior,* 24, 595-605.

Rose, J., Jenkins, R., O'Connor, C., Jones, C. & Felce, D. (2002). A group treatment for men with intellectual disabilities who sexually offend or abuse. *Journal of Applied Research in Intellectual Disabilities,* 15, 138-50.

**Saum, S. W. (2007). A comparison of an actuarial risk prediction measure (static-99) and a stable dynamic risk prediction measure (stable-2000) in making risk predictions for a group of sexual offenders. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 68*(3-B), 1943.

*Seager, J. A., Jellicoe, D., & Dhaliwal, G. K. (2004). Refusers, dropouts, and completers: Measuring sex offender treatment efficacy. *International Journal of Offender Therapy and Comparative Criminology, 48,* 600-612.

Segal, Z.V., & Marshall, W.K. (1985). Heterosexual social skills in a population of rapists and child molesters. *Journal of Consulting and Clinical Psychology, 53,* 55-63.

Segal, Z.V., & Marshall, W.K. (1986). Discrepancies between self-efficacy predictions and actual performance in a population rapists and child molesters. *Cognitive Therapy and Research, 10,* 363-376.

Seidman, B.T., Marshall, W.L., Hudson, S.M., & Robertson, P.J. (1994). An examination of intimacy and loneliness in sex offenders. *Journal of Interpersonal Violence, 9,* 518-534.

*Seto, M. C. (2005). Is more better? combining actuarial risk scales to predict recidivism among adult sex offenders. *Psychological Assessment, 17,* 156-167.

Seto, M. C., Harris, G. T., Rice, M. E., & Barbaree, H. E. (2004). The screening scale for pedophilic interests predicts recidivism among adult sex offenders with child victims. *Archives of Sexual Behavior, 33,* 455-466.

Sjöstedt, G., & Grann, M. (2002). Risk assessment: What is being predicted by actuarial prediction instruments? *International Journal of Forensic Mental Health, 1,* 179-183.

Sjöstedt, G., & Långström, N. (2001). Actuarial assessment of sex offender recidivism risk: A cross-validation of the RRASOR and the static-99 in Sweden. *Law and Human Behavior, 25,* 629-645.

Sjöstedt, G., & Långström, N. (2002). Assessment of risk for criminal recidivism among rapists: A comparison of four different measures. *Psychology, Crime and Law, 8,* 25-40.

*Skelton, A., Riley, D., Wales, D., & Vess, J. (2006). Assessing risk for sexual offenders in New Zealand: Development and validation of a computer-scored risk measure. *Journal of Sexual Aggression, 12,* 277-286.

*Soothill, K., Harman, J., Francis, B., & Kirby, S. (2005). Identifying future repeat danger from sexual offenders against children: A focus on those convicted and those strongly suspected of such crime. *Journal of Forensic Psychiatry and Psychology, 16,* 225-247.

Sreenivasan, S., Garrick, T., Norris, R., Cusworth-Walker, S., Weinberger, L. E., Essres, G., Turner, S., & Fain, T. (2007). Predicting the likelihood of future sexual recidivism: Pilot study findings from a California sex offender risk project and cross-validation of the static-99. *Journal of the American Academy of Psychiatry and the Law, 35,* 454-468.

Sroufe, L.A., & Rutter, M. (1984). The domain of developmental psychopathology. *Child Development, 55, 17*-29.

Stadtland, C., Hollweg, M., Kleindienst, N., Dietl, J., Reich, U., & Nedopil, N. (2005). Risk assessment and prediction of violent and sexual recidivism in sex

offenders: Long-term predictive validity of four risk assessment instruments. *Journal of Forensic Psychiatry and Psychology, 16*, 92-108.

Stadtland, C., Hollweg, M., Kleindienst, N., Dietl, J., Reich, U., & Nedopil, N. (2006). Evaluation of risk assessment instruments for sex offenders. *Nervenarzt, 77*, 587-595.

*Sudo, J., Sato, M., Obata, S., & Yamagami, A. (2006). Exploring the possibility of risk assessment of Japanese sexual offenders using Static-99. *Criminal Behaviour and Mental Health, 16*, 146-154.

Swanson, C.K. & Garwick, B. (1990). Treatment for low-functioning sex offenders: Group therapy and interagency coordination. *Mental Retardation*, 28, 155-161.

Thompson, D. & Brown, H. (1997). Men with intellectual disabilities who sexually abuse: a review of the literature. *Journal of Applied Research in Intellectual Disabilities,* 10, 140-158.

Thornton, D. (1995). *Self-esteem Questionnaire*. (Unpublished).

Thornton, D. (2002). Constructing and testing a framework for dynamic risk assessment. *Sexual Abuse: A Journal of Research and Treatment,* 14, 139-154.

Thompson, D., & Browne, H. (1997). Men with intellectual disabilities who sexually abuse: A review of the literature. *Journal of Applied Research in Intellectual Disabilities, 10*, 140-158.

Thornton, D., Mann, R., Webster, S., Blud, L., Travers, R., Friendship, C., & Erikson, M. (2003). Distinguishing and combining risks for sexual and violent recidivism. In R. Prenky, E. Janus, M. Seto, and A.W. Burgess (Eds), *Understanding and managing sexually coercive behavior. Annals of the New York Academy of Sciences, 989*, 225-235.

Torgerson, C.J. (2006). Publication bias: the achilles' heel of systematic reviews? *British Journal of Educational Studies, 54*, 89-102.

Tough, S.E (2001). *Validation of two standardised risk assessments (RRASOR, 1997: Static-99, 1999) on a sample of adult males who are developmentally disabled with significant cognitive deficits.* Unpublished Master's Thesis, University of Toronto, Toronto, Ontario, Canada.

Tudway, J., & Darmoody, M. (2005). Clinical assessment of adult sexual offenders with learning disabilities. *Journal of Sexual Aggression, 11,* 277-288.

*Vrieze, S. I., & Grove, W. M. (2008). Predicting sex offender recidivism. I. correcting for item overselection and accuracy overestimation in scale development. II. sampling error-induced attenuation of predictive validity over base rate information. *Law and Human Behavior, 32,* 266-278.

*Ward, T., & Beech, A. R. (2004). The etiology of risk: A preliminary model. *Sexual Abuse: A Journal of Research and Treatment, 16,* 271-284.

Ward, T., Polaschek, D.L.L., & Beech, A.R. (2006).*Theories of sexual offending*. Chichester: Wiley.

*Watson, T., & Vess, J. (2007). Risk assessment of child-victim sex offenders for extended supervision in New Zealand. *Journal of Forensic Psychiatry and Psychology, 18,* 235-247.

Webster, C. D., Harris, G. T., Rice, M. E., Cormier, C., & Quinsey, V. L. (1994). *The violence predictions scheme*. Toronto, Ontario: Centre for Criminology, University of Toronto.

*Webster, S.D., Mann, R.E., Carter, A.J., Long, J., Milber, R.J., O'Brien, M.D., Wakeling, H.C., & Ray, N.L. (2006). Inter-rater reliability of dynamic risk assessment with sexual offenders. *Psychology, Crime and Law, 12,* 439-452.

Webster, S.D., Mann, R.E., Thornton, D., & Wakeling, H.C. (2007). Further validation of the short self-esteem scale with sexual offenders. *Legal and Criminological Psychology, 12,* 207-216.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised.* New York: Psychological Corporation.

Wechsler, D. (1997). *Weschsler Adult Intelligence Scale-III*. San Antonio, TX: The Psychological Corporation.

Wilcox, D.T. (2004). Treatment of intellectually disabled individuals who have committed sexual offences: A review of the literature. *Journal of Sexual Aggression,* 10, 85-100.

Wilcox, D., Beech, A. R., Markall, H.F., & Blacker, J. (2009). Actuarial risk assessment and recidivism in a sample of U.K. intellectually disabled sex offenders. *Journal of Sexual Aggression, 15,* 97-106.

Williams, F., Wakeling, H., & Webster, S. (2007). A psychometric study of six self-report measures for use with sexual offenders with cognitive and social functioning deficits. *Psychology, Crime & Law, 13,* 505-522.

Witte, T. D., Di Placido, C., Gu, D., & Wong, S. C. P. (2006). An investigation of the validity and reliability of the criminal sentiments scale in a sample of treated sex offenders. *Sexual Abuse: A Journal of Research and Treatment, 18,* 249-258.

Worling, J., & Cruwen, T. (2000). Adolescent sexual offender recidivism: success of specialized treatment and implications for prediction. *Child Abuse & Neglect, 24, 965-982.*

*Yates, P. M., & Kingston, D. A. (2006). The self-regulation model of sexual offending: The relationship between offence pathways and static and dynamic sexual offence risk. *Sexual Abuse: A Journal of Research and Treatment, 18,* 259-270.

Zachary, R.A. (1996). *Shipley Institute of Living Scale: Revised Manual.* Los Angeles, CA: Western Psychological Services.

*Zanatta, R. G. (2006). Risk of violent and sexual recidivism: A comparison of dangerous offenders and repetitive sexual offenders. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 67*(3-B), 1723.

APPENDICES SECTION

Appendix 1

Search Syntax

PsycINFO on OVID search syntax searched 14 July 2008, period 1985 to week 2 July 2008

1. exp Sex Offenses/ (21219)
2. exp sexual abuse/ or incest/ or exp rape/ (17769)
3. paraphilias/ or exhibitionism/ or fetishism/ or incest/ or pedophilia/ or sexual masochism/ or sexual sadism/ or voyeurism/ (4528)
4. child molest$.mp. [mp=title, abstract, heading word, table of contents, key concepts] (759)
5. (risk adj4 (assess$ or evaluat$)).mp. [mp=title, abstract, heading word, table of contents, key concepts] (10423)
6. risk management/ (1801)
7. risk management.mp. [mp=title, abstract, heading word, table of contents, key concepts] (2477)
8. risk assessment.mp. [mp=title, abstract, heading word, table of contents, key concepts] (6078)
9. exp Risk Assessment/ (4457)
10. predictive validity.mp. (4229)
11. exp Statistical Validity/ (11550)
12. exp Measurement/ or exp Prediction/ or exp Risk Assessment/ or exp Recidivism/ or exp Statistical Analysis/ (200562)
13. 1 or 2 or 3 or 4 (23035)
14. 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 (207774)
15. 13 and 14 (2639)

MEDLINE on OVID search syntax searched 11 July 2008, period 1996 to week 1 July 2008

1. exp Sex Offenses/ (7329)
2. exp sex offenses/ or child abuse, sexual/ or rape/ (7329)
3. exp sex offenses/ or exp child abuse, sexual/ or exp rape/ (7329)
4. exp paraphilias/ or exp exhibitionism/ or exp "fetishism (psychiatric)"/ or exp masochism/ or exp pedophilia/ or exp sadism/ or exp transvestism/ or exp voyeurism/ (866)
5. exp Risk Assessment/ (82026)
6. exp Risk Management/ (96918)
7. (risk adj4 assess$).mp. [mp=title, original title, abstract, name of substance word, subject heading word] (96139)
8. statistical validity.mp. [mp=title, original title, abstract, name of substance word, subject heading word] (76)
9. "Sensitivity and Specificity"/ (158308)
10. actuarial analysis/ or area under curve/ or confidence intervals/ or "sensitivity and specificity"/ (187447)
11. 1 or 2 or 3 or 4 (7887)
12. 5 or 6 or 7 or 8 or 9 or 10 (289491)
13. 11 and 12 (489)

14. from 13 keep 1-10 (10)
15. from 13 keep 1-385 (385)

EMBASE on OVID search syntax searched 11 July 2008, period 1988 to week 28 2008

1. exp Sexual Crime/ or sex offences.mp. (5456)
2. sexual abuse/ (7116)
3. sexual abuse.mp. or exp Sexual Abuse/ (13465)
4. sexual crime/ (3379)
5. incest/ or rape/ (2849)
6. exp Sexual Deviation/ (2127)
7. paraphilias.mp. (185)
8. exhibitionism/ or fetishism/ or masochism/ or pedophilia/ or sadism/ (1246)
9. child molest$.mp. (176)
10. risk assessment/ (166930)
11. risk assessment.mp. or Risk Assessment/ (170101)
12. Risk Management/ (10149)
13. actuarial.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name] (11081)
14. predictive validity/ (1000)
15. predictive validity.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name] (2482)
16. (risk adj4 (assess$ or evaluat$)).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name] (186760)
17. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 (15014)
18. 10 or 11 or 12 or 13 or 14 or 15 or 16 (206279)
19. 17 and 18 (984)
20. from 19 keep 1-984 (984)

ISI Web of Knowledge – (Science Citation Index, Social Science Citation Index Arts and Humanities Citation Index) on ISI search syntax searched 11 July 2008, period 1996 to 2007

1 Topic=("sex offen*") 5,366
2 Topic=("sex* abuse*") 13,717
3 Topic=(incest) 3,950
4 Topic=(rape) 22,255
5 Topic=(exhibitionis*) 639
6 Topic=(pedophil*) 991
7 Topic=("sexual sadism") 47
8 Topic=("risk assessment") >100,000
9 Topic=("risk management") 22,950
10 Topic=("predictive validity") 4,195
11 #10 OR #9 OR #8 >100,000
12 #7 OR #6 OR #5 OR #4 OR #3 OR #2 OR #1 41,985
13 #12 AND #11 921

(ASSIA) - CSA searched syntax searched 17 July 2008

Query: (((sex* abuse) or (sex* offen*) or (child molest*)) or (rape or (sexual sadism) or exhibitioni*) or pedophil*) and(((risk assessment) or (risk management) or (predictive validity)) or (statistical validity)) or (actuarial analysis)) or ((area under curve) or (sensitivity or specificity)))

Appendix 2

Inclusion/Exclusion Form

Author:
Year:
Title:
Source:

Inclusion criteria	Criteria met?	Comments
Population: over 17 years old AND been conviction for a sexual offence	YES NO	
Intervention: Has an actuarial risk tool been completed? (specifically for sexual offenders)	YES NO	
Outcome: How is recidivism measured? • Reconviction • Re-arrest • Cautions • Self-report		
Study type: • cohort • case control (retrospective or prospective)		
Exclusion: Narratives, commentaries, editorials or other types of opinion paper	YES NO	

Appendix 3

Quality Assessment Form A - Cohort

Author(s):
Year:
Title:
Source:

Question	Score			U	Comments
	Y(2)	N(0)	P(1)		
SELECTION BIAS					
Was the study objectives clear?					
Was the cohort recruited in an acceptable way?					
Was the selection of participant's representative of the defined population? (Was there something special about the cohort?)					
MEASUREMENT/DETECTION BIAS					
Was there a clear definition of sexual recidivism?					
Was the method of outcome assessment clearly stated? (Reconviction data sources identified Was outcome measure clearly identified?)					
Were the outcome assessors blind? (Did they know who were the recidivists and non-recidivists)					
Was the outcome measured in the same way across all participants?					
Was the risk assessment tool administered by trained individuals?					
Were the risk assessments scored using multiple sources of information? (Please circle- archival files, psychometric tests, interviews with offenders, other)					
Was the inter-rater reliability above the .8 threshold?					
Was there a sufficient follow-up for the outcome to occur? (2 years minimum)					
Was missing information dealt with appropriately? (i.e. missing info needed to complete risk assessments)					
ATTRITION BIAS					
Was the actual rate of participant drop-out recorded?					
Did the study record the stage at which participants dropped out?					

RESULTS					
What are the results? (Was ROC reported, sensitivity/specificity and/or likelihood ratios presented?)					
Was the predictive validity of the test/s clearly stated?					
Was concurrent validity of the test/s addressed or discussed?					
Was the standardisation of the test/s described?					
Do the results of the study fit with other available evidence?					
Are the results of the study reliable? (Are the design/ methods of the study sufficiently flawed to make the results unreliable?)					
Can the results be generalised to other populations? (Consider age, ethnicity, offender type)					

Quality Score: / **No. of Unclear:**

Appendix 4

Quality Assessment Form B - Case Control

Author(s):
Year:
Title:
Source:

Question	Score			U	Comments
	Y(2)	N(0)	P(1)		
SELECTION BIAS					
Was the study objective clear?					
Were the cases representative of the target population?					
Were the controls reliably assessed?					
Were the cases reliably assessed?					
Were the controls clearly defined					
Were the cases clearly defined?					
Was the control/comparison group randomly selected from the source population?					
Is the description of the demographic factors of the population clear?					
Were the cases and controls comparable in terms of demographic/confounding variables?					
MEASUREMENT/DETECTION BIAS					
Was there a clear definition of sexual recidivism?					
Was the method of outcome assessment clearly stated? (Reconviction data sources identified Was outcome measure clearly identified?)					
Was the outcome measured in the same way across cases and controls?					
Was the risk assessment tool administered by trained individuals?					
Were the risk assessments scored using multiple sources of information? (Please circle- archival files, psychometric tests, interviews with offenders, other)					
Was the inter-rater reliability above the .8 threshold?					
Was the follow-up period long enough for the outcome to occur? (2 yrs minimum)					
Was missing information dealt with					

appropriately? (i.e. missing info needed to complete risk assessments)					
ATTRITION BIAS					
Were the dropout rates and reasons similar in the case and control groups?					
Did the study record the stage at which participants dropped out?					
RESULTS					
What are the results? (Was ROC reported, sensitivity/specificity and/or likelihood ratios presented?)					
Was the predictive validity of the test/s clearly stated?					
Was concurrent validity of the test/s addressed or discussed?					
Do the results of the study fit with other available evidence?					
Are the results of the study reliable? (Are the design/ methods of the study sufficiently flawed to make the results unreliable?)					
Can the results be generalised to other populations? (Consider age, ethnicity, offender type)					
Were the potential confounding factors of the study design and/or analysis taken into account?					

Quality Score: / **No. of Unclear:**

Appendix 5

Data Extraction Form

Author (s):

Year:

Title:

Sources:

Location of study:

Type of study:

Name of risk assessment tool:

Total Sample Size:

Risk category:

Number of High risk:

Number of Low risk:

Re-offence rate:

Number reoffended:

Number non-reoffended:

Outcome measure:

Length of follow-up period:

Statistical Analysis/Results:

ROC analysis:

Likelihood Ratio:

Sensitivity:

Specificity:

Inter-rater reliability:

Quality assessment score: **Clarity score:**

Appendix 6

Excluded Studies and Reasons for Exclusion

Study	Reasons for Exclusion
Austin, Rayton & Johnson (2003)	Did not meet the threshold criteria based on clear description of outcome measure
Barbaree, H. E., Langton, C. M., & Peacock, E. J. (2006a)	Did not meet the threshold criteria based on clear description of outcome measure
Barbaree, H. E., Langton, C. M., & Peacock, E. J. (2006b)	No outcome measure
Bates et al. (2004)	Did not meet the threshold criteria based on sufficient statistical analysis
Beech, A., & Ford, H. (2006)	Did not meet the threshold criteria based on sufficient statistical analysis
Berlin, Galbreath, Geary & McGlone (2003)	Narrative
Boer, Tough & Haaven (2004)	Commentary
Campbell (2003)	Commentary
Craig, Browne & Stringer (2003)	Narrative
Craig, Browne & Stringer (2004)	No outcome measure
Craig, Browne, Stringer & Hoguee (2008)	Review paper
Craissati & Beech (2006)	Intervention not appropriate. Risk tools were not used in isolation. Other variables were combined into tool.
Craissati, Falla, McClurg & Beech (2002)	Did not meet the threshold criteria based on sufficient statistical analysis
Dempster & Hart (2002)	Intervention not appropriate. Uses SVR-20 items not the overall tools
Doren (2004)	Did not meet the threshold criteria based on sufficient statistical analysis
Doren (2006)	Commentary
Ferguson (1998)	No outcome measure
Ferguson, Eidelson & Witt (1998)	No outcome measure
Furr (1993)	Intervention not appropriate
Gentry, Dulmus & Theriot (2005)	No outcome measure
Green, Gray & Willner (2003)	No outcome measure
Hanson, Broom & Stephenson (2004)	Intervention not appropriate. Risk factor items from two instruments were used.
Hanson & Bussière (1998)	Intervention not appropriate. No risk instruments used
Hanson & Harris (2000)	Intervention not appropriate. Used SONAR a measure of need not risk
Hanson & Morton-Bourgon (2005)	Intervention and outcome measures not appropriate
Harris & Tough (2004)	Did not meet threshold criteria based on clearly defined outcome measure
Johnson (2003)	No outcome measure
Langevin (2006)	Outcome measure not appropriate.

	Specifically used sexual homicide
Levenson et al (2006)	Did not meet threshold criteria based on clearly defined outcome measure
Levenson (2004)	No outcome measure
Lindsay, Hogue, Taylor, Mooney, Steptoe, Johnston et al. (2006)	No outcome measure
Lindsay, Hogue, Taylor, Steptoe, Mooney, O'Brien, et al. (2008)	No outcome measure
Lindsay, Murphy, Smith, Murphy, Edwards, Chittock, et al. (2004)	Intervention and population not appropriate. Sexual and violent offenders were not differentiated in analysis
Mossman (2006)	Commentary
Nunes & Cortoni (2008)	No outcome measure
Nunes, Firestone et al. (2007)	Intervention not appropriate (modified version of RRASOR). Also did not meet the threshold criteria based on sufficient statistical analysis
Nunes, Hanson et al. (2007)	Intervention not appropriate (modified version of RRASOR). Also did not meet the threshold criteria based on sufficient statistical analysis
Proulx, Pellerin, Paradis, McKibben, Aubut & Ouimet (1997)	Intervention not appropriate. Used risk factors rather than risk instrument
Rice & Harris (1997)	Intervention not appropriate. VRAG was used.
Roberts, Doren & Thornton (2002)	No outcome measure
Seager, Jellicoe & Dhaliwal (2004)	Did not meet the threshold criteria based on sufficient statistical analysis
Seto (2005)	Intervention not appropriate. Risk tools were combined and not used in isolation
Skelton, Riley, Wales & Vess, (2006)	Intervention not appropriate. Computer scored risk instrument
Soothill, Harman, Francis & Kirby (2005)	Did not meet the threshold criteria based on sufficient statistical analysis
Sudo, Sato, Obata & Yamagami (2006)	No outcome measure
Vrieze & Grove (2008)	Did not meet the threshold criteria of a clear description of the outcome measure
Ward & Beech (2004)	Narrative
Watson & Vess (2007)	No outcome measure
Webster, Mann, Carter, Long, Milner, O'Brien, et al. (2006)	Intervention not appropriate and not outcome measure
Yates & Kingston (2006)	No outcome measure
Zanatta (2006)	No outcome measure

Appendix 7

Table: Characteristics of cohort studies reporting the predictive validity of risk assessment instruments

Authors, year and location	**Sample**	**Risk tools used**	**Average Duration of outcome**	**Outcome measure**	**Inter-rater reliability**	**Sexual Re-offence rate**	**Predictive Accuracy Findings (including AUC, *r*, SE & CI when detailed)**[1]	**Quality (no. unclear)**
Hanson & Thornton (2000) UK and Canada	1301 mixed sex offenders from 4 samples, prison and psychiatric settings	RRASOR Static-99 SACJ-Min	Averaged ranged 4-23 years based on 4 samples	Convictions or Charges	No details reported	15.4%-35.1%	*RRASOR* AUC=.68, *r* =.28 (95%CI=.65-.72) *Static-99* AUC=.71, *r* =.33 (95%CI=.68-.74) *SACJ-Min* AUC=.67, *r* =.23 (95%CI=.63-.71)	21/40 (7/20)
Barbaree, Seto et al (2001) Canada	215 rapists and child molesters from prison setting	SORAG RRASOR Static-99 MnSOST-R	4.5yrs (SD=2.2, range=29 days to 9.9yrs)	Convictions or Charges	No details reported	9%	*SORAG* AUC= .70 (*r*=.70 p<.001) *RRASOR* AUC=.77(*r*=.77 p<.001) *Static-99* AUC=.70 (*r*=.70 p<.001) *MnSOST-R* AUC=. 65 (*r*=.65, n.s)	30/40 (4/20)
Sjöstedt & Långström	1400 rapists and child	RRASOR Static-99	3.96 yrs (SD=	Convictions	Average *k* =.90	4%	*RRASOR* AUC=.72, *r*=.22 p<.01, (95%CI=.64-	31/40

(2001) Sweden	molesters from a prison setting		1.39, range= 1 to 6 yrs				.80) *Static-99* AUC=. 76, *r*= 22 p<.01, (95%CI=.65-.72)	(1/20)
Beech, Friendship, Erikson & Hanson (2002) UK	140 child abusers from a community/ probation setting	Static-99	6 yrs	Convictions	No details reported	15% (8/53)	*Static-99* AUC= .77, *r*= 31 p<.05, (95% CI=. 55-.98)	22/40 (3/20)
English, Retzlaff & Kleinsasser (2002) USA	494 sex offender (218 on probation, 47 on parole, 229 in prison)	CO-SOMB	12 or 30 mths	Conviction, negative treatment termination, commission of new sex crime (no details of type of offences)	No details reported	54% (n=267 at 12 mths) 40% (n=197 at 30 mths)	*CO-SOMB* AUC=.64 The odds ratio of failing was higher for offenders with high risk scale scores (Odds ratio of 2.05, 95% confidence band=1.21 to 3.47)	13/40 (7/20)
Nunes et al. (2002) Canada	258 sex offenders (138 incest, 107 extra familial child molesters, and 13	Modified Static-99 SORAG	7.3 yrs (SD= 3.5, range= 0.8 to 25 yrs)	Convictions or Charges	No details reported	8.9% (n=23)	*Static-99* AUC= .70, SE= .05 (95% CI=. 66-.79) *SORAG* AUC= .65, SE= .06 (95% CI=. 52-.98)	25/40 (4/20)

	rapists). From a hospital setting							
Rice & Harris (2002) Canada	184 offenders (82 daughter molesters, 102 extra familiar molesters) From a mixture of probation/ community and hospital settings	SORAG	53.6 mths	Re-arrests	.98 Pearsons correlation	Not clear	*SORAG* AUC= .65 –father daughter offenders AUC= .81, SE= .05, (95% CI=± .09) – all offenders *r*= .42 p<.001	23/40 (5/20)
Sjöstedt & Långström (2002) Sweden	51 rapists (all with personality disorder, from prison, probation and hospital settings)	RRASOR SVR-20	92.33 mths (SD= 31.33, range= .80 to 136.4 mths	Convictions	Cohen's k's were .60 or higher ICC= .62 or higher	20% (n=10)	*RRASOR* AUC= .73, (95% CI=. 56-.90), *r*= .10 n.s *SVR-20* AUC= .49, (95% CI=. 26-.71), *r*= .10 n.s	25/40 (4/20)
Sjöstedt & Grann (2002) Sweden	1303 sex offenders (both rapists & child molesters)	RRASOR Static-99	6 yrs (SD= 1.39, range= 3 to 8 yrs)	Reoffending Imminence (reoffence within 1	No details reported	0.2% (within a month) 3%	*RRASOR* AUC= .73, (95% CI=. 67-.88) AUC= .92, (95% CI=. 85-.99) - imminence AUC= .75, (95% CI=. 65-.86) -	24/40 (7/20)

	From prison setting			mth) & Frequency (repeated reoffending)		repeated	repeated reoffending *Static-99* AUC= .75, (95% CI=. 68-.81) AUC= .94, (95% CI=. 87-1.02) - imminence AUC= .75, (95% CI=. 65-.85) - repeated reoffending	
Thornton (2002) UK	117 offenders (detailed offence-type info not reported) From prison setting	Static-99	5.8 yrs	Convictions	No details reported	n =7	AUC= .92, SE= .035 (95% CI=. 85-.99)	17/40 (8/20)
Bartosh, Garby, Lewis & Gray (2003) US	186 mixed (70 rapists, 44 intra-familiar, 53 extra-familiar and 19 non-contact) From prison setting	RRASOR Static-99 MnSOST-R SORAG	60-66 mths	Sexual recidivism (not stated)	No details reported	11.8% (n=22)	*RRASOR* AUC= .632, *r*= .154 p<.05 (all offenders) AUC=. 575(extra-fam), AUC=.727(intra-fam), AUC= .534 (rape), AUC= .492 (hands-off) *Static-99* AUC= .636 , *r*= .140 n.s (all offenders) AUC=. 647(extra-fam), AUC=.735(intra-fam), AUC= .714 (rape), AUC= .394 (hands-off) *MnSOST-R* AUC= .585, *r*= .096 n.s (all offenders)	26/40 (5/20)

							AUC=. 586(extra-fam), AUC=.629 (intra-fam), AUC= .536 (rape), AUC= .568 (hands-off) *SORAG* AUC= .579, *r*= .072 n.s (all offenders) AUC=. 697(extra-fam), AUC=.723(intra-fam), AUC= .707 (rape), AUC= .477 (hands-off)	
Harris & Rice (2003) Canada	Kingston & Pacific sample (no sex offender details, including numbers or demographic information) One sample was from a prison setting	RRASOR Static-99	Fixed follow-up	Sexual recidivism – no details	Kingston sample – r's & kappa's >.80 Other sample had no inter-rater reliability	No details given	*RRASOR* AUC=. 71 (sample with interrater reliability assessed) at 2yr follow-up AUC=.70 (sample with no interrater reliability) at 2yr follow-up AUC=. 60 (sample with interrater reliability assessed) at 11yr follow-up AUC=.55 (sample with no interrater reliability) at 11yr follow-up *Static-99* AUC=. 80 (sample with interrater reliability assessed) at 2yr follow-up AUC=.72 (sample with no interrater reliability) at 2yr follow-up AUC=. 59 (sample with interrater reliability assessed) at 11yr follow-up AUC=.59 (sample with no interrater reliability) at 11yr follow-up In the sample were interrater reliability was assessed the AUC was greater. The	18/40 (7/20)

							longer the follow-up period the smaller the AUC became.	
Harris, Rice et al (2003) Canada	396 from 4 samples (rapists and child molesters) From community and prison settings	RRASOR Static-99 SORAG	61.5 mths (SD= 48.1)	Convictions or Charges	RRASOR (ICC=.95) Static-99 (ICC=.87) SORAG (ICC=.96)	26% (n= 104)	*RRASOR* AUC= .59, SE= .03 (95% CI=. 52-.65) AUC=.61 (child molesters) AUC=.56 (rapists) *Static-99* AUC= .62, SE= .03 (95% CI=. 56-.68) AUC=.65 (child molesters) AUC=.59 (rapists) *SORAG* AUC= .66, SE= .03 (95% CI=. 60-.71) AUC=.70 (child molesters) AUC=.62 (rapists)	34/40 (2/20)
Thornton et al. (2003) UK	1910 untreated sexual offenders (sample 1 contained 647 & sample 2 contained 429) From a prison setting	RM2000/S	3.1yrs (sample 1) & 19yrs (sample 2)	Conviction	No details reported	2.6% (n= 17) Sample 1 27.7% (n= 119) Sample 2	AUC= .77 (sample 1 treated offenders) AUC= .75 (sample 2 untreated offenders)	19/40 (7/20)

Craig, Brown, Beech & Stringer (2004) UK	121 sexual and violent offenders (63 sex offenders, non-mentally disordered) From a prison setting	Static-99 RM2000/S	8yrs (SD=9.9, range 5yrs 6mths to 10 yrs 3 mths)	Conviction	No details reported	16.2% (n=11) at 2 yrs 26.5% (n=18) at 5 yrs 34.4% (n=22) at 10yrs	*Static-99* AUC=.59 (2yr follow-up) AUC=.58 (5yr follow-up) AUC=.52 (10yr follow-up) *RM2000* AUC=.56 (2yr follow-up) AUC=.58 (5yr follow-up) AUC=.55 (10yr follow-up)	27/40 (3/20)
deVogel, de Ruiter, Van Beek & Mead (2004) Netherlands	122 sex offenders (95 rapists & 26 extra familiar child molesters & 1 incest) From a hospital setting	Static-99 SVR-20	140 mths	Conviction	Static-99 ICC=.80 SVR-20 (total) ICC=. 75	39%	*Static-99* AUC=. 71, SE=.05, (*r*=.38 p<.001) *SVR-20* AUC=. 80, SE=.04, (*r*=.50 p<.001)	29/40 (2/20)
Långström (2004) Sweden	1303 mixed sex offenders, divided into 3 ethnic groups (95% rapists and child molesters) From a	RRASOR Static-99	5.7yrs (SD= 1.4, range 3.0 to 8.0 yrs)	Conviction	No details reported	No details reported	*RRASOR* AUC=.73 (95%CI=.66-.79) AUC=.76, *r*=.30 p<.01 (Nordic group) AUC=.77, *r*=.52 p<.01 (European group) AUC=.48, *r*=.02 n.s (African group) *Static-99* AUC=.75 (95%CI=.68-.81)	34/40 (0/20)

	prison setting						AUC=.76, *r*=.27 p<.01, (Nordic group) AUC=.79, *r*=.42 p<.01, (European group) AUC=.50, *r*=.01 n.s (African group)	
Seto, Harris, Rice & Barbaree (2004) Canada	258 (two samples of 115 & 145) all child molesters From a prison setting	RRASOR Static-99 SORAG	5.0yrs (study 1) 64 mths (study 2)	Charge	ICC's ranged from .90 to.94	8% (n= 9) Study 1 19% (n= 25) Study 2	Study 1 *RRASOR* AUC= .83, SE= .05, (*r*=.55, p<.001) *Static-99* AUC= .81, SE= .05, (*r*=.44, p<.001) *SORAG* AUC= .74, SE= .09, (*r*=.32 p<.001) Study 2 *RRASOR* AUC= .69, SE= .06, (*r*=.61 p<.001) *Static-99* AUC= .72, SE= .06, (*r*=.54 p<.001) *SORAG* AUC= .74, SE= .06, (*r*=.27 p<.005)	25/40 (4/20)
Craissati & Beech (2005)	310 offenders (80 rapists, 230 child molesters) From community/ probation setting	Static-99 RM2000	No details given	Conviction and probation breach	4 yrs (only for 273 offenders)	2%	*Static-99* AUC= .713 (total sample) AUC= .530 (rapists) AUC= .768 (child molesters) *RM2000* AUC= .70 (total sample) AUC= .667 (rapists) AUC= .713 (child molesters)	26/40 (4/20)

							RM2000 (χ^2=17.842, df3, p<.01) and Static-99 (χ^2=17.767, df3, p<.01) were significantly associated with failure.	
Firestone et al. (2005) Canada	656 mixed offenders (hands-on offenders with adult or child victims) From a hospital setting	RRASOR-Mod	12.18 yrs (SD=3.33)	Conviction or Charges	No details reported	19.7%	No AUC's performed *r*(637)= .22 P<.001	20/40 (5/20)
Stadtland et al. (2005) German	134 offenders; 3 samples – (73 therapeutic prison sample; 15 treatment dropout sample; 46 assessment sample, accused of sex offence)	Static-99 SVR-20	9yrs (range 1 to 340 mths)	Conviction	No details reported	27.6% (n=37) Contact sex offence 9% (n=12) non-contact Sex offence	*Static-99* AUC=.742 (All non-contact sexual) AUC=.662 (All contact sexual) *SVR-20* AUC=.544 (All non-contact sexual) AUC=.684 (All contact sexual)	32/40 (0/20)

Allan, Dawson & Allan (2006) Australia	538 offenders (347 Indigenous & 191 non-Indigenous). From a prison setting	RRASOR Static-99	9yrs 3 mths	Conviction	No details reported	n= 179 non-violent sexual n= 69 violent sexual offence	*RRASOR* AUC=.71 (all offenders) AUC=.74 (Non-Indigenous) AUC=.65 (Indigenous) For Indigenous offenders the classification accuracy was poor for reoffenders (17.1%). For non-Indigenous classification accuracy was 33.6% for reoffenders. *Static-99* AUC=.78 (Non-Indigenous nonviolent sexual, n=144). Classification accuracy was moderate for reoffenders (52.5%) and good for non-reoffenders (91.3%).	19/40 (7/20)
Craig, Browne, Beech & Stringer (2006a) UK	85 sex offenders (all contact offenders) Sample from a hospital setting. No offenders with psychotic illness	SVR-20	8yrs 7mths (SD= 9yrs 5mths, range from 5yrs 6mths to 10yrs 3 mths)	Conviction	No details reported	19% (2yrs) 28% (5yrs) 36% (10 yrs)	AUC=.48, *r*=-.01 n.s (2yr follow up) AUC=.48, *r*=-.02 n.s (5yr follow up) AUC=.48, *r*=-.02 n.s (10yr follow up)	22/40 (6/20)

Craig, Browne, Beech & Stringer (2006b) UK	119 offenders (66 from a regional secure unit, mostly child abusers & 53 child abusers from the community)	Static-99	72 mths & 106 mths	Conviction	No details reported	6% (n=7) at 2yrs 12% (n=14) at 5yrs	AUC=.67 (95%CI=.47-.86) – 2yr follow up AUC=.61 (95%CI=.43-.77) – 5yr follow up AUC=.62 (95%CI=.47-.76) – 10yr follow up	21/40 (7/20)
Craig, Beech & Browne (2006) UK	85 contact sexual offenders & 46 non-sexual violent offenders. Sample from a hospital setting. No offenders with a mental disorder.	RM2000-S Static-99 SVR-20	8yrs & 7mths (SD=9.5mths, range from 5yrs 6mths to 10 yrs 3mths	Conviction	No details reported	7% (n=6) of sex offender group	For sex offender group only *RM2000-S* AUC=.60, *r*=.05 n.s (2yr follow up) AUC=.68, *r*=.08 n.s (5yr follow up) AUC=.59, *r*=.05 n.s(10yr follow up) *Static-99* AUC=.57, *r*=.01 n.s (2yr follow up) AUC=.59, *r*=.02 n.s (5yr follow up) AUC=.52, *r*=.00 n.s (10yr follow up) *SVR-20* AUC=.46, *r*=-.03 n.s (2yr follow up) AUC=.48, *r*=-.02 n.s (5yr follow up) AUC=.51, *r*=-.00 n.s (10yr follow up)	25/40 (4/20)
Ducro & Pham (2006) Belgium	147 offenders from a high security hospital (63.8% child	Static-99 SORAG	4.2 yrs (SD= 3.4 yrs)	Sexual recidivism and/or repeated return to a	Static-99 ICC= .63 SORAG ICC=. 92	25.2%	*Static-99* AUC=.66, SE=.05, *r*=.23 p<.01 (95%CI=.56-.76) – Total sample AUC=.70, SE=.06, *r*=.33 p<.01 (95%CI=.59-.82) – Child abusers	28/40 (4/20)

	abusers, 24.6% rapists, 11.5% mixed)			forensic unit			AUC=.71, SE=.10, *r*=.15 n.s (95%CI=.56-.76) – Rapists *SORAG* AUC=.64, SE=.06, *r*=.18 p<.05 (95%CI=.53-.75) – Total sample AUC=.65, SE=.07, *r*=.19 p<.01 (95%CI=.52-.78) – Child abusers AUC=.64, SE=.14, *r*=.21 n.s (95%CI=.36-.92) – Rapists	
Hanson (2006) Canada, USA & UK	3424 sex offenders from 8 samples combined. Mixture of prison, hospital and community settings	Static-99	Varied average of 2 to 23yrs depending on sample	Conviction or Charge	ICC=.91 and *K*=.87	15.7%	AUC=.68 (95%CI=.62-.74) – age18-24 AUC=.68 (95%CI=.64-.72) – age 25-39 AUC=.66 (95%CI=.58-.73) – age 40-49 AUC=.76(95%CI=.66-.85) – age 50-59 AUC=.82 (95%CI=.68-.95) – age 60+ AUC=.70 (95%CI=.67-.72) – All ages	23/40 (6/20)
Looman (2006) Canada	258 treated high risk mixed sex offenders. From a hospital setting	Static-99 SORAG	5.1 yrs	Conviction	r= .90 (Static-99) r=.91 (SORAG)	8.9% (n=23)	*Static-99* AUC=.63 *SORAG* AUC=.69	34/40 (1/20)

Stadtland et al. (2006) Germany	134 sex offenders (details not known)	Static-99 SVR-20	9yrs (range from 1 to 340 mths)	Conviction	No details reported	27.6%	*Static-99* AUC=.721, SE=.050 (95%CI=.624-.818) – Excluding treatment dropouts AUC=.710, SE=.045 (95%CI=.621-.799) – Including treatment dropouts SVR-20 (total score) AUC=.682, SD=.05 (95%CI=.584-.780) Including treatment dropouts does not improve predictive accuracy	*N/A due to translation*
Witte, Di Placido, Gu &Wong (2006) Canada	72 offenders (33 rapists, 20 pedophiles, 10 incest & 9 mixed). From a hospital setting	Static-99	3yrs (54.5mths, SD= 9.7, range from 39.5 to75.8 mths)	Conviction	ICC=.94 (n=10)	18.3%	AUC= .72, *r*=.29 p<.05	20/40 (6/20)
Allen et al. (2007) New Zealand	495 (47.7% extra-familial, 52.3% intra-familial victims). From a prison setting	Static-99	5.8yrs	Conviction	No details reported	9.9% (n=49)	*Static-99* AUC= .72, *r*= .27 (95% CI= .64-.80) *Static-99 combined with deviance* AUC= .81, (95% CI= .76-.87)	24/40 (5/20)

Bengtson & Långström (2007) Denmark	121 (60 rapists, 51 extra-familial & 8 extra-familial). From a hospital setting	Static-99 Static-2002	16.3yrs (SD= 4.0, range from 7.7 to 24.4yrs	Conviction	*k*=.86 (Static-2002) *k*=.89 (Static-99)	31%	*Static-99* AUC= .62, (95% CI= .52-.72) AUC= .72, (95% CI= .59-.84) –severe sexual recidivism *Static-2002* AUC= .67, (95% CI= .57-.77) AUC= .69 (95% CI= .56-.83) –severe sexual recidivism Sensitivity scores of .40 (Static-99) and .47 (Static-2002) Specificity scores of .82 (Static-99) and .72 (Static-2000)	37/40 (0/20)
Craig, Thornton, Beech & Browne (2007) UK	119 child molesters. From a community and hospital assessment setting.	Static-99	6 yrs	Conviction	No details reported	6% (n=7) at 2 yrs 12% (n=14) at 5 yrs	AUC= .66, *r*= .007 n.s - at 2yrs AUC= .60, *r*= .114 n.s - at 5yrs	20/40 (7/20)
Dietiker, Dittman & Graf (2007) Switzerland	64 sex offenders (details not known) From a hospital setting.	SVR-20	Not clear	Conviction	No details reported	Not clear	SVR-20 AUC= .88 Sensitivity=.85 Specificity=77 Cut-off of 12.5	*N/A due to translation*

Knight & Thornton (2007) USA	599 (266 sexually dangerous sex offenders; & 333 non-sexually dangerous sex offenders). From a prison setting	SORAG Static-99 Static-2002 RM2000 SVR-20 MnSOST-R	No mean follow-up stated. Fixed follow-up periods of 5 and 10 yrs	Conviction & Charges	Ranged from .602 to .887. Only SVR-20 was below.70	21.5% at 15yrs (total sample)	*SORAG* AUC=.67 (95%CI=.60-.74) at 5yr AUC=.64 (95%CI=.54-.74) at10yr *Static-99* AUC=.68 (95%CI=.62-.75) at 5yr AUC=.65 (95%CI=.56-.74) at 10yr *Static-2002* AUC=.67 (95%CI=.61-.74) at 5yr AUC=.67 (95%CI=.58-.76) at 10yr *RM2000-S* AUC=.64 (95%CI=.57-.71) at 5yr AUC=.63 (95%CI=.54-.73) at 10yr *SVR-20* AUC=.68 (95%CI=.62-.75) at5yr AUC=.68 (95%CI=.59-.77) at 10yr *MnSOST-R* AUC=.67 (95%CI=.60-.74) at 5yr AUC=.66 (95%CI=.56-.76) at 10yr	36/40 (0/20)
Langton, Barbaree, Seto, Peacock, Harkins & Hanson (2007)	468 offenders (175 rapists, 155 child molesters, 93 familial & 45 mixed). From a	RRASOR Static-99 Static-2002 SORAG MnSOST-R	5.9yrs (SD= 3.0yr, range from 1 day to 11.7 yrs	Conviction	ICC's ranged from .88 to.94	11%	*RRASOR* AUC= .68, (95% CI=. 61-.75) *Static-99* AUC= .64, (95% CI=. 57-.71) *Static-2002*	38/40 (1/20)

Canada	prison setting						AUC= .71, (95% CI=. 64-.78) *SORAG* AUC= .66, (95% CI=. 58-.74) *MnSOST-R* AUC= .70, (95% CI=. 62-.77)	
Langton, Barbaree, Hanson Harkins & Peacock (2007) Canada	464 (32 refused treatment, 389 completed treatment & 38 dropped out). From a prison setting	RRASOR, Static-99 Static-2002	5.9 yrs	Conviction	α=.68 (Static-2002 only)	No details reported	*RRASOR* AUC= .67, (95% CI=. 59-.75) - Completers AUC= .71, (95% CI=. 48-.94) - Refusers AUC= .82, (95% CI=. 69-.95) – Dropouts *Static-99* AUC= .61, (95% CI=. 53-.68) - Completers AUC= .75, (95% CI=. 55-.75) - Refusers AUC= .82 (95% CI=. 65-.99) – Dropouts *Static-2002* AUC= .69, (95% CI=. 61-.77) - Completers AUC= .84, (95% CI=. 63-.1.00) - Refusers AUC= .81, (95% CI=. 61-.1.00) –	28/40 (5/20)

							Dropouts Static-2002 had a likelihood ratio of 2.25 at 5yrs, with a recidivism score of 7. Sensitivity and specificity cut-offs were also presented, a score of 5 or less was considered low risk (score of 7+ high risk).	
Olver et al. (2007) Canada	321 offenders High intensity SOTP (169 rapists, 56 child molesters, 45 mixed, 51 incest). From a hospital setting	Static-99 VRS-SO	10.0yrs (SD=4.0, range from 2.0 to 19 yrs)	Conviction	Static-99 ICC=.82 VRS-SO α=.84	24.6% (n=79)	*Static-99* AUC= .63, (95% CI=. 56-.70) *VRS-SO* AUC= .74, (95% CI=. 68-.80)	35/40 (2/20)
Screenivasen et al. (2007) USA	137 sex offenders (both contact and non-contact offenders). From a prison setting	Static-99	13.80 yrs (SD= .85)	Conviction, arrests, parole/probation violations	No details reported	31% - 5yrs 40%-10yrs	AUC=. 62, *r*= .24, p<.01	27/40 (4/20)

Bengtson (2008) Denmark	304 (160 rapists, 144 child molesters; 37 intra-familial & 107 extra-familial All psychiatric assessed with mental illness. From a psychiatric setting.	Static-2002 Static-99 RM2000/S	16.2 yrs (SD=4.3, ranged from 1.96 to 24.42 yrs)	Conviction	Static-02 *k*= .86 ICC=.96 *Static-99* *k*= .89 ICC= .94 *RM2000/S* *k*=.85 ICC= .72	28%	*Static-99* AUC=.64, SE=.03 (total) AUC=.64, SE=.05 (rapists) AUC=.67, SE=.05 (child molesters) *Static-2002* AUC=.67, SE=.03 (total) AUC=.68, SE=.05 (rapists) AUC=.69, SE=.05 (child molesters) *RM2000* AUC=.65, SE=.03 (total) AUC=.61, SE=.05 (rapists) AUC=.71, SE=.05 (child molesters)	30/40 (3/20)

[1] AUC = Area under the curve as measured by receiver operating characteristic analyses, r = Pearson's correlation, SE = Standard error, CI = Confidence interval, ICC= Intraclass correlation coefficient, *K*= Kappa, α= cronbach alpha.

Table: Characteristics of case control studies reporting the predictive validity of risk assessment instruments

Authors, year and location	**Sample**	**Risk tools used**	**Duration of outcome**	**Sexual Recidivism Outcome measure**	**Inter-rater reliability**	**Sexual Re-offence rate**	**Findings and Statistical analysis (including AUC, SE & CI when detailed)**[1]	**Quality (no. unclear)**
Hanlon, Larson & Zacher (1999)	26 sex offenders (10 were reincarcerated, 10 were not reincarerated)	MnSOST	38 to 62 mths	Conviction (excluded new untried legal charges)	No details reported	N=4	Correlation between MnSOST and recidivism scores yielded a Pearson correlation (*r*) of 0.482 and p= .016 (one-tailed).	25/52 (5/26)
Hanson & Harris (2001) Canada	409 offenders (non-incestuous hands-on). From a community setting	Static-99 SONAR	24 mths	Convictions & Charges	94%-97% average agreement	n= 208	SONAR AUC= .74, (*r*=.43, P<.01) Static-99 no AUC reported. Recidivists were at higher risk than non-recidivists on Static-99 (r=.15, p<.01).	40/52 (4/26)
Tough (2001) Canada	76 intellectually disabled sex offenders. From a community setting	Static-99 RRASOR (modified scoring to reduce recidivism info being lost)	8.58 yrs recidivists 5.52 yrs non-recidivists	Convictions, re-arrests & Verbal reports by community members, family or agency staff	No details reported	n= 12 15.8%	RRASOR correlated significantly with recidivism (r=3.05, p<.05) and was able to able to distinguish between recidivists and non-recidivists (t= 2.752, p<.05) Static-99 did not significantly correlate with recidivism (r=.080, n.s) and was not able to distinguish recidivists and non-recidivists (t = 0.691, n.s) No ROC curve analysis completed	26/52 (5/26)

[1] AUC = Area under the curve as measured by receiver operating characteristic analyses, r = Pearson's correlation, SE = Standard error, CI = Confidence

Appendix 8

Name:……. Date:…….
DOB:…….
Ethnicity:……
Age at time of treatment:……
Name of Probation Officer:……

SVR-20 Coding Sheet

Psychosocial Adjustment	Presence (N, ?, Y)	Recent change (+, 0, -)
1. Sexual deviation		
2. Victim of Child abuse		
3. Psychopathy		
4. Major mental illness		
5. Substance use problems		
6. Suicidal/homicidal ideation		
7. Relationship problems		
8. Employment problems		
9. Past nonsexual violent offences		
10. Past non-violent offences		
11. Past supervision problems		
Sexual Offences		
12. High Density sex offences		
13. Multiple sex offence types		
14. Physical harm to victims(s) in sex offences		
15. Uses weapons or threats of death in sexual offences		
16. Escalation in frequency or severity of sex offences		
17. Extreme minimisation or denial of sex offences		
18. Attitudes that support or condone sex offences		
Future Plans		
19. Lacks realistic plans		
20. Negative attitudes towards intervention		
Other Considerations		

Summary of risk rating	
1. Risk of sexual violence	Low/Moderate/High

Appendix 9

RISK MATRIX 2000

Family Name: Forenames: DOB:
Case Id: Assessor: Date of Assessment:

RM2000/S – Risk of Sexual Offending

STEP ONE

1) Age at Commencement of Risk

	Points
Under 18	0
18-24	2
25-34	1
Older	0

2) Sexual Appearances

	Points
1	0
2	1
3,4	2
5 or more	3

3) Criminal Appearance

	Points
Less than 5	0
5 or more	1

Total Points	Category
0	Low risk
1,2	Medium risk
3,4	High risk
5,6	Very high risk

STEP TWO: AGGRAVATING FACTORS

A) Any convictions for a contact sex offence against a male Yes No
(do not count consenting sexual behaviour involving only males aged 16 years or older)

B) Any conviction for a contact sex offence against a stranger Yes No
(count as stranger of the victim did not know the offender 24 hours before the offence)

C) "Single" Yes No
(Count as "single" if never lived with an adult lover for at least 2 years)

D) Any conviction for a non-contact sex offence? Yes No

Put up one risk category if two aggravating factors are present and up two categories if four aggravating factors are present

STATIC RISK CLASSIFICATION – RISK OF SEXUAL OFFENDING

Low Risk **Medium Risk** **High Risk** **Very High Risk**

RM2000/V Risk of Violent Offending

1) Age at Commencement of Risk	Points
Under 18	4
18 to 24	3
25 to 34	2
35 to 44	1
Older	0

2) Sexual Appearances	Points
0	0
1	1
2,3	2
4 or more	3

3) Any Burglaries?	Points
No	0
Yes	2

Total Points	Category
0,1	Low risk
2,3	Medium risk
4,5	High risk
6+	Very high risk

STATIC RISK CLASSIFICATION – RISK OF VIOLENT OFFENDING

Low Risk **Medium Risk** **High Risk** **Very High Risk**

Appendix 10

The Rapid Risk Assessment for Sexual Offense Recidivism (RRASOR).

Prior sex offenses (not including index offenses)	
None	0
1 conviction; 1-2 charges	1
2-3 convictions; 3-5 charges	2
4 or more convictions; 6 or more charges	3
Age at release (current age)	
more than 25	0
than 25	1
Victim gender	
only females	0
any males	1
Relationship to victim	
only related	0
any non-related	1

Appendix 11

ARMIDILLO[1] SCORING SHEET[2]

Name: _____ Age: _____

Specify time period for evaluating recent change: _____

RISK FACTOR CATEGORIES:

Stable Dynamic Items (Environmental Factors)	**Presence -2 to +2**	**Recent Change (+, 0, -)**	**Critical Item?**
1. Attitude towards intellectually, learning or developmentally disabled individuals			
2. Communication among supervisory staff			
3. Client specific knowledge by supervisory staff			
4. Consistency of supervision			
5. Situational consistency			
6. Generalisation of treatment concepts			
TOTAL SUB-SECTION SCORE:			
Acute Dynamic Items (Environmental Factors)	**Presence -2 to +2**	**Recent Change (+, 0, -)**	**Critical Item?**
1. Changes in social relationships			
2. New supervisory staff			
3. Monitoring of offender by staff			
4. Victim access			
5. Situational changes			
6. Access to intoxicants			
TOTAL SUB-SECTION SCORE:			
TOTAL SECTION SCORE:			

[1] The ARMIDILLO is a structured risk and management guideline instrument under development by Boer, Haaven, Lambrick, Lindsay, McVilly, Sakdalan, and Smith. It is intended for use with intellectually, developmentally, or learning disabled individuals (youth or adults) for whom there are concerns regarding violent or sexually violent behaviour which may or may not have been adjudicated.

[2] Draft date: July 21, 2007

Stable Dynamic Items (Client)	**Presence -2 to +2**	**Recent Change (+, 0, -)**	**Critical Item?**
1. Attitude toward and compliance with supervision			
2. Attitude toward and compliance with treatment			
3. Knowledge of faulty thoughts, behaviour cycle, risk factors and relapse prevention plan			
4. Sexual knowledge and self-management of sexuality			
5. Mental health problems, self-knowledge, monitoring ability and self-management			
6. Time management skills and planning ability			
7. Substance abuse			
8. Victim selection and acquisition / grooming behaviour			
9. General coping ability and self-efficacy			
10. Relationship skills			
11. Use of violence or threats of violence towards self or others			
12. Impulsiveness			
TOTAL SUB-SECTION SCORE:			
Acute Dynamic Items (Client)	**Presence -2 to +2**	**Recent Change (+, 0, -)**	**Critical Item?**
1. Changes in substance abuse pattern			
2. Changes in sexual preoccupation			
3. Changes in emotional state and / or changes in ability to manage emotional changes			
4. Changes in victim access or related behaviours			
5. Changes in attitude or behaviour toward supervision or treatment			
6. Changes in ability to use coping strategies			
TOTAL SUB-SECTION SCORE:			
TOTAL SECTION SCORE:			

Summary of Risk Rating: **Low** **Moderate** **High**

Assessment completed by: _____

Date of assessment: _____

Appendix 12

RAS questionnaire

Directions: Indicate how characteristic or descriptive each of the following statements is of you by using the coded given

+3	+2	+1	-1	-2	-3
Very like me	Rather like me	Somewhat like me	Somewhat unlike me	Rather unlike me	Very unlike me

1	Most people seem to be more aggressive and assertive than I am	
2	I have hesitated to make or accept dates because of 'shyness'	
3	When the food is served at a restaurant is not done tom satisfaction, I complain about it to the waiter	
4	I am careful to avoid hurting other people's feelings, even when I feel that I have been injured	
5	If a salesman has gone to considerable trouble to show me merchandise which is not quite suitable, I have a difficult time saying 'no'	
6	When I am asked to do something, I insist upon knowing why.	
7	There are times when I look for a good vigorous argument.	
8	I strive to get ahead as well as most people in my position	
9	To be honest, people often take advantage of me	
10	I enjoy staring conversations with new people	
11	I often don't know what to say to attractive persons of the opposite sex	
12	I will hesitate to make phone calls to business establishment and institutions	
13	I would rather apply for a job by writing letters than going through interviews	
14	I find it embarrassing to return products	
15	If a close relative were annoying me, I would smother my feelings rather than express my annoyance	
16	I have avoided asking questions for fear of sounding stupid.	
17	During an argument I am sometimes afraid that I will get so upset that I will shake all over.	
18	If someone I respect makes a statement which I think is incorrect, I will have them hear my point of view as well.	
19	I avoid arguing over prices with shop assistants	
20	When I have done something importance or worthwhile, I manage to let others know about it.	
21	I am open and frank about my feelings	
22	If someone has been spreading false stories about me, I will speak to them about it as soon as possible.	
23	I often have a hard time saying 'no'.	
24	I tend to bottle up my emotions rather than make a scene	
25	I complain about poor service in a restaurant and elsewhere.	
26	When I am given a compliment, I sometimes just don't know what to	

	say.	
27	If a couple near me in a cinema were talking loudly, I would ask them to be quiet	
28	Anyone attempting to push ahead of me in a line is in for a good battle.	
29	I am quick to express my opinion.	
30	There are times when I just can't say anything.	

Adapted Self-esteem questionnaire

1. Do you wish you were someone else?.....Yes No
2. Do you like the sort of person you are?.....Yes No
3. Do you often feel ashamed of yourself?.....Yes No
4. Do you understand yourself?.....Yes No
5. Do you think you can make a success of your life?.....Yes No
6. Are things all mixed up in your life?..... Yes No
7. Are you pretty happy with the way you are?.....Yes No
8. Do you have a low opinion of yourself?.....Yes No

Appendix 13

Consent Form

I have been approached by a student of Forensic Psychology from the University of Birmingham.

Janine Blacker has stated that she would like to conduct a case study about me as part of her course work. This case study will consist of an account of my historical background, the details of my offence, my treatment and the results of that treatment.

I understand that I will not be identified in any written work and that my confidentiality is assured.

I consent to information recorded at Probation being used in the aforementioned case study.

I have been informed that I have the right to withdraw my consent at any time.

I an aware that although my name will appear o this form, it will not be reproduced in any written documents which form part of the case study and that I will be referred to by a false name which bears no resemblance to my own name.

Signed:

Signature witnessed by: Signed:

Date:

If the participant wished to received further information, it is to be sent to the address below in completed confidentiality.

Copy of the consent form supplied to participants: YES/NO

Signed:

Appendix 14

Case Study Worksheets and Materials

Eye Contact Record

SESSION NUMBER: 1

DATE: 12/03/03

Each time I make eye contact during the session- an eye lash can be drawn on

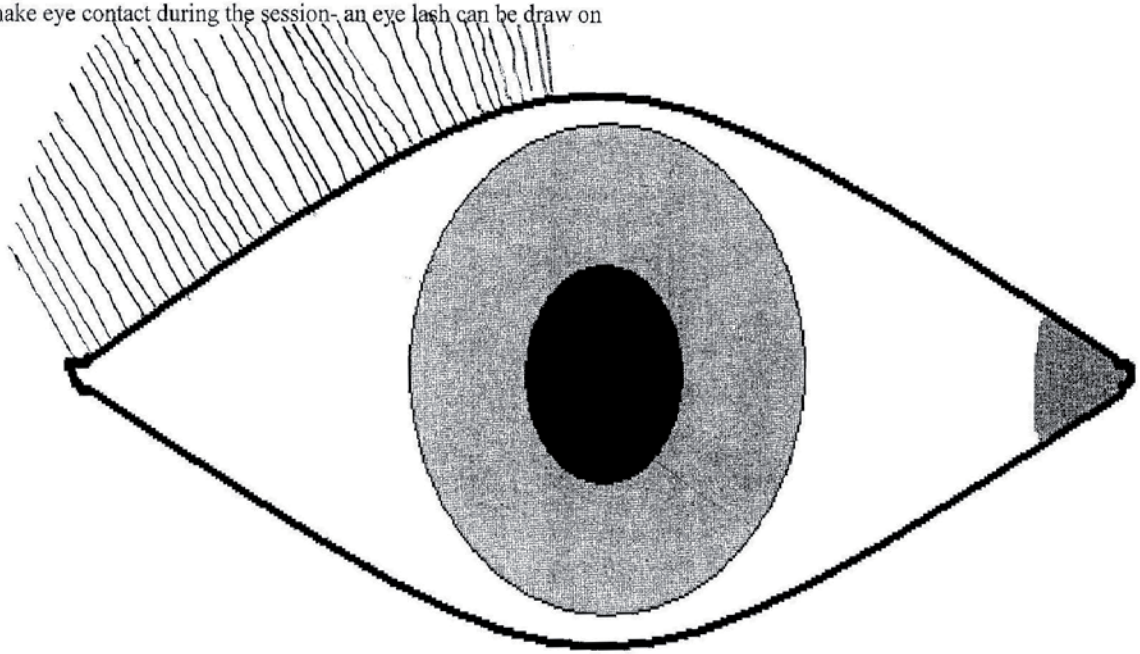
14

Eye Contact Record

SESSION NUMBER: 10

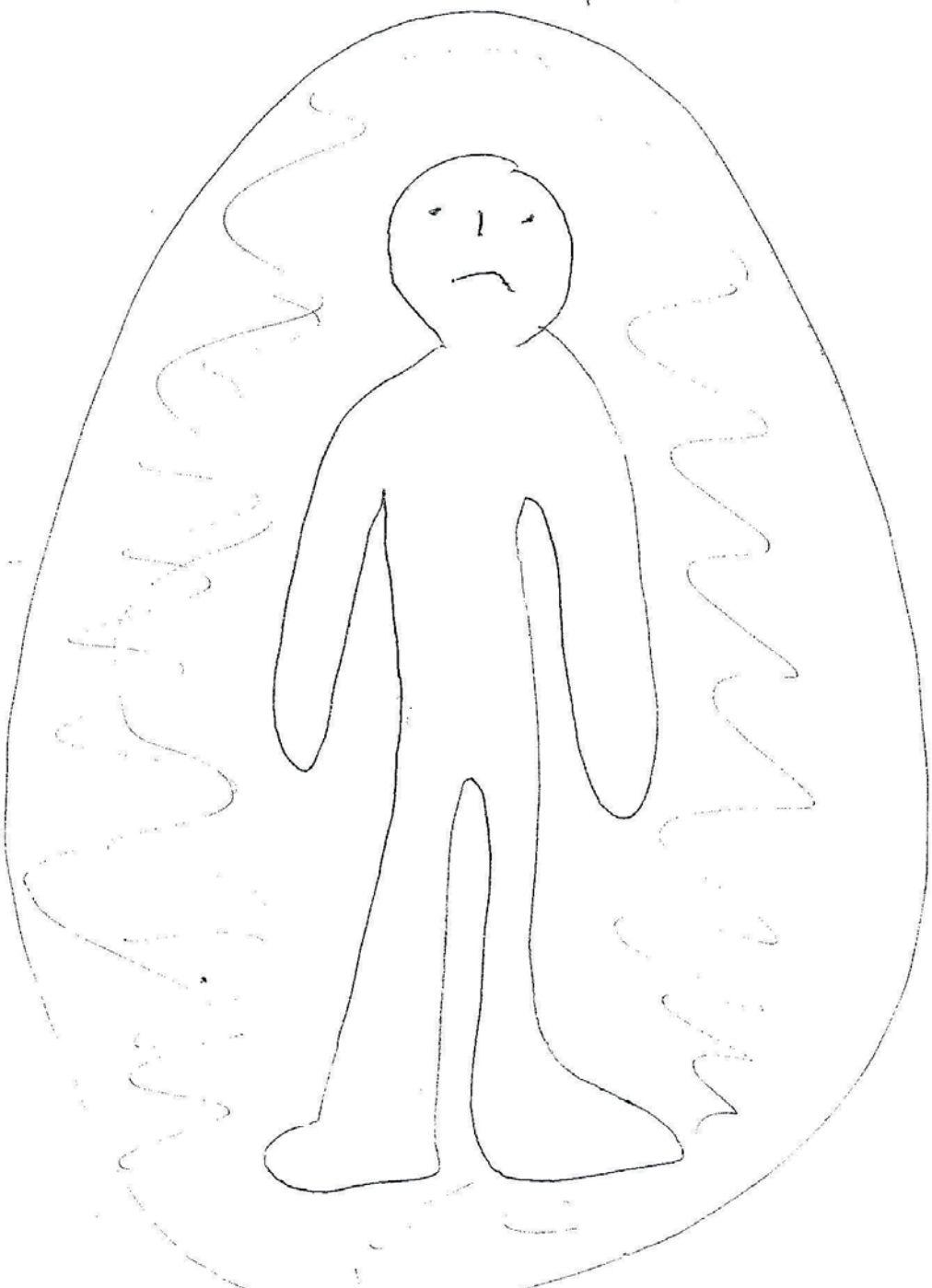
DATE: 12/04/07

Each time I make eye contact during the session- an eye lash can be draw on



SMALLER SPACE = friends, Relatives, girlfriend

LARGER SPACE = Landlord, people I dont know
probation, police



uncomfortable, unsure what other person will say.

→ not listening
→ people treated me badly — told me off
→ taken advantage of me

3 Behaviours

Previously mumbled a lot

	Aggressive	Passive	Assertive
Voice	Shouting	Whining Mumbling under their breath	Calm
Body Language	Folded arms Pointing finger	Shuffling feet Downcast eye Stoop / Slouching	Direct eye contact Upright posture
Words used	You'd better…. If you don't Should	Maybe I guess Sorry sorry sorry but	I I think I feel I want How can we resolve this

Consequences

I feel you need to help me with the bill
I think we need to pay the bills equally to make it fairer

(1)

How did I FEEL in the ~~past week~~?

when I got bullied

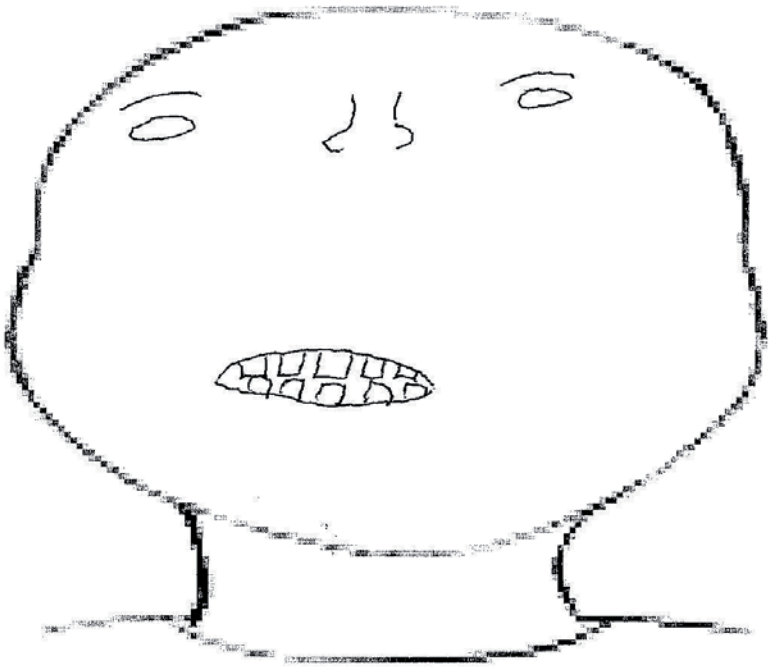
WHY? feel really down & like I am being tortured by them.
They are doing something they 'should' not do

⟹ feeling lonely

FEELING USED

(2)

How did I FEEL in the ~~past week~~? when I get bullied

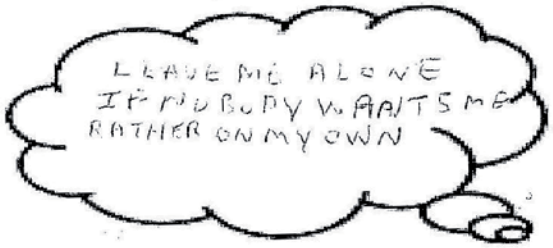


WHY? AGGRESSION
I LIKE THEM TO KEEP AWAY
I LIKE MY OWN SPACE ZONE

BULLYING

OLD ME

Thinking

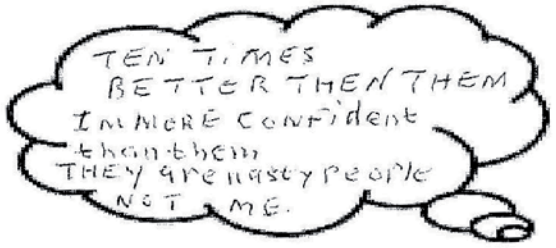


Feeling

Result LONELY
STOP TALKING
To People about
MYSELF

NEW ME

Thinking



Feeling

Result
GONE COPE WITH
IT BETTER
GO OUT MAKE NEW FRIENDS

How did I FEEL ~~in the past week~~?

flatmate not paying bills.

WHY? FELT TIMID

SCARED WHAT HE GONNA SAY —

I AM NOT GOING TO PAY

THE BILL

BEFORE

SCRATCH MY HEAD

AND LOOK FLUFFY

MUMBLING

New Me

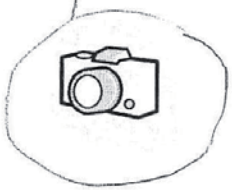
EYE CONTACT

STAY CALM

TELL CLEARLY

EXPLAIN QUICKLY

get back into photography



LEISURE TIME

1. Darts, snooker, pool
   (club in dudley (flatmates friends))

2. Rabbling Club
   (Rotary Club do walks)

3. Social Club

4. Watching football

5. Photography
   (When retire - join the British legion Club)

go to Silverstone or Donnington to watch racing with flatmate

THINK GOOD – FEEL GOOD

My relaxing activities

Fill in the thought bubbles by writing or drawing the things that help you to relax and feel calm.

THINK GOOD – FEEL GOOD

What happens when I feel happy?

Think about something that made you feel really happy.
How would someone else know that you felt like this?

What does your face look like when you are happy?

How does your body show that you are happy?

JUMP FOR JOY

How do you behave when you are happy?

GO OUT SOMEWHERE NICE
PROUD

How much of the time do you feel happy?

Never 1 2 3 4 (5) 6 7 8 9 10 All the time

THINK GOOD – FEEL GOOD

What happens when I feel angry?

Think about something that made you feel really cross and angry. How would someone else know that you felt like this?

What does your face look like when you are angry?

How does your body show that you are angry?

CLENCHED FIST

How do you behave when you are angry?

BASH THE TABLE
KICK THE DOOR

How much of the time do you feel angry?

Never
1 2 3 4 5 6 (7) 8 9 10
All the time

Small Steps

At times it might be useful to break down the task into smaller steps

Each step feels more manageable

This increases the chances of success and each step will move you closer to your target

Your Task: MAKE A FRIEND

- ✓ Step 1 WENT TO LIBRARY TO ASK ABOUT CLUBS
- Step 2 CALLING THEM UP ABOUT SPORTING EVENTS
- Step 3 WHAT BUS TO CATCH
- Step 4 GOING TO THE PLACES
- Step 5 MEET A STRANGER

STEP6 MEETING OTHER PEOPLE
STEP7 GIVE TELEPHONE NUMBER
STEP8 GOING PLACES TOGETHER

Appendix 15

Client's Diary

DIARY SHEET NUMBER:.....

DATE	PLACE/SITUATION	WHO WAS THE PERSON	WHAT DID WE TALK ABOUT	HOW SUCCESSFUL WAS I?
21/03/07	Sobriehill Library	Two female librarians	Voluntary work Smoker clubs I asked for information on social clubs.	Confident going in on at ostinly for info Eye contact could have been better
22/03/07	City Centre	Old friend from hotel I use to work at	We asked how I was doing and was I working. I showed him my camera on my phone & took his photo	Felt fine Full eye contact not made. Head down looked at him for five seconds
20/03/07	City Centre	Market Researcher	Walking towards the person. She said have you got any time, I said no I haven't got time and walked passed her.	Very successful as I said no.

DIARY SHEET NUMBER:.....

DATE	PLACE/SITUATION	WHO WAS THE PERSON	WHAT DID WE TALK ABOUT	HOW SUCCESSFUL WAS I
28 WED MARCH 2007	ON THE ROAD	MIDDLE AGE WOMAN	ABOUT BILLS AND MY MOBILE AND HER FRIEND WAS A LITTLE STRANGE TO ME	I WAITED TILL SHE WENT UP THE ROAD. I DEALT WITH IT OK
WED 28TH 2007	STAYED IN ~~RE~~ AT HOME	MYSELF	GOT UP LATE FOR THE COURCE. THEY CALLED.	I DEALT WITH IT OK. I WAS SAD BECAUSE IT WAS MY [illegible]
			A FEW DAYS BEFORE I SIGNED FOR IT. AND I TOLD THEM ABOUT ME	WRONG DOING IN FIRST PLACE

DIARY SHEET NUMBER:.....

DATE	PLACE/SITUATION	WHO WAS THE PERSON	WHAT DID WE TALK ABOUT	HOW SUCCESSFUL WAS I
13/03/07	on bus after the session	An old lady	She was going home to make some muffins. Asked where I can from. Waved goodbye.	felt ok. Was speaking politely Practice eye contact
18/3/07	PAPER SHOP	CHILDREN IN SHOP	THEY TALKED ABOUT EASTER WHEN I S COMING. I WALKED ROUND THEM	FELT OK. BUT I FELT I DEALT WITH IT OK.
16/3/07	On a road to a shop	Neighbour 21 yr old woman (known her 4 years)	She asked if I had an appointment with an agency	felt I had good eye contact. She kept staring at me. I kept my distance physically.
15/3/07	On the bus	Jamaican lady	I said to her I am going home to make my tea & said to her I like rice & peas. She replied I like rice & peas too. She got off the bus & smiled.	felt strange ~~to~~ to make a conversation with a Jamaican lady. She didn't make much eye contact & looked away.

DIARY SHEET NUMBER;.....

DATE 1	PLACE/SITUATION	WHO WAS THE PERSON	WHAT DID WE TALK ABOUT	HOW SUCCESSFUL WAS I
1ST APRIL 2007	AT HOME IN GARDEN	TWO NEIGH BOURS	ABOUT DOING THEIR GARDEN FOR THEM	A BIT QUICK EYE CONTACT
WED [illegible]	IN SOLIHULL	SEMI RETIRED LADY	TALKING ABOUT RETIREMED	FULL ON EYE CONTACT IT WAS UK ASK MORE
TUES 5TH	IN MY ROAD	OLD LADY	TALK ABOUT PAYING THE BILLS	QUESTION EYE CONTACT